

## **Ética 4.0: Dilemas morais nos cuidados de saúde mediados por robôs sociais**

Antonio Alvaro Soares

Mestrado em Ciências em Emoções

Orientadora:

Professora Doutora Patrícia Paula Lourenço Arriaga Ferreira, Professora Auxiliar,  
ISCTE – Instituto Universitário de Lisboa

Coorientador:

Doutor Nuno José Guerreiro Piçarra,  
ISCTE – Instituto Universitário de Lisboa

Setembro, 2021

## **Ética 4.0: Dilemas morais nos cuidados de saúde mediados por robôs sociais**

Antonio Alvaro Soares

Mestrado em Ciências em Emoções

Orientadora:

Professora Doutora Patrícia Paula Lourenço Arriaga Ferreira, Professora Auxiliar,  
ISCTE – Instituto Universitário de Lisboa

Coorientador:

Doutor Nuno José Guerreiro Piçarra,  
ISCTE – Instituto Universitário de Lisboa

Setembro, 2021

À Lara Gesteira Bäuerlein, com amor.



## **Agradecimentos**

Agradeço à Lara Bäuerlein pela paciência e compreensão perante as intermináveis horas em que fiquei recluso a pesquisar, escrever, fazer cálculos e refletir sobre esta investigação.

Agradeço à Patrícia Arriaga e ao Nuno Piçarra pelas orientações e pelo constante convite à reflexão e estudo.

Agradeço de modo especial aos mais de 840 participantes que voluntariamente cederam parte de seu tempo e responderam ao nosso inquérito.



## Resumo

A Inteligência Artificial e os robôs sociais nos cuidados de saúde trazem um novo campo de investigação interdisciplinar. Neste estudo examinámos os julgamentos morais das pessoas acerca da reação de uma agente de saúde perante uma paciente que recusa uma medicação. Para o efeito, desenvolvemos um dilema moral que variou em função da agente (humana vs. robô), decisão (respeito à autonomia vs. beneficência/não-maleficência) e argumentação (benefício à saúde vs. prejuízo à saúde). Avaliámos a aceitabilidade moral da decisão, a responsabilidade moral da agente e os seus traços de amabilidade, competência e confiabilidade, atribuídos por 524 participantes (350 mulheres; 316 brasileiros, 179 portugueses; 18-77 anos), aleatorizados por 8 vinhetas, num desenho inter-sujeitos aplicado através de um inquérito *online*. Os julgamentos de aceitabilidade moral foram mais elevados na decisão do respeito à autonomia da paciente, evidência similar para as duas agentes. A responsabilização moral e a perceção de amabilidade foram superiores para a humana em relação à robô. Não houve diferenças na perceção de competência e confiabilidade das agentes. As agentes que respeitaram a autonomia foram percebidas como muito mais amáveis, com uma dimensão de efeito superior aos outros atributos, mas menos competentes e confiáveis que as agentes que decidiram pela beneficência/não-maleficência. As agentes que priorizaram a beneficência/não-maleficência e argumentaram acerca do benefício à saúde foram consideradas mais confiáveis do que nas demais interações entre a decisão e a argumentação. Esta investigação contribui para a compreensão dos julgamentos morais no contexto dos cuidados de saúde mediados por agentes tanto humanos como artificiais.

Palavras-chave: Psicologia moral da robótica, julgamentos morais, cuidados de saúde, robôs de enfermagem, amabilidade & competência

Classificação nas categorias e códigos definidas pela American Psychological Association (PsycINFO): 2300 Psicologia Experimental Humana; 2340 Processos Cognitivos; 2360 Motivação e Emoção; 3360 Psicologia da Saúde e Medicina; 3365 Promoção e Manutenção da Saúde e Bem-estar; 3370 Serviços de Saúde e Saúde Mental; 3375 Cuidados Domésticos e Hospitais; 3377 Enfermagem e Cuidados Residenciais; 3400 Questões Profissionais do Pessoal da Psicologia e da Saúde; 3430 Atitudes e Características do Pessoal Profissional; 3450 Ética Profissional, Normas e Responsabilidade; 4100 Sistemas Inteligentes; 4120 Inteligência Artificial e Sistemas Especializados; 4140 Robótica.





## Abstract

Artificial Intelligence and social robots in healthcare bring a new interdisciplinary field of research. In this study, we have examined people's moral judgments about a healthcare agent's reaction to a patient who refuses a medication. For this purpose, we have developed a moral dilemma that was varied according to the type of healthcare agent (human vs. robot), decision (respect for autonomy vs. beneficence/non-maleficence), and argumentation (health benefit vs. health harm). We have assessed the decision's moral acceptability, the agent's moral responsibility, and her traits of warmth, competence, and trustworthiness assigned by 524 participants (350 women; 316 Brazilian, 179 Portuguese; 18-77 years old) randomized into 8 vignettes, in an inter-subject design that was applied using an online survey. Moral acceptability judgments were higher in the decision to respect patient autonomy, similar evidence for both agents. Moral responsibility and perceived warmth were higher for the human agent than for the robot, and there were no differences in the agents' perceived competence and trustworthiness. Agents who have respected autonomy were perceived as much warmer, with a higher effect dimension than the other attributes, but less competent and trustworthy than agents who have decided for beneficence/non-maleficence. Agents who have prioritized beneficence/non-maleficence and argued about the health benefit were perceived as more trustworthy than in the other interactions between decision and argumentation. This research contributes to the understanding of moral judgments in the context of healthcare mediated by both humans and artificial agents.

Keywords: Moral psychology of robotics, moral judgments, healthcare, nursing robots, warmth & competence

Classification as defined by the American Psychological Association (PsycINFO Classification Categories and Codes): 2300 Human Experimental Psychology; 2340 Cognitive Processes; 2360 Motivation & Emotion; 3360 Health Psychology & Medicine; 3365 Promotion & Maintenance of Health & Wellness; 3370 Health & Mental Health Services; 3375 Home Care & Hospice; 3377 Nursing Homes & Residential Care; 3400 Professional Psychological & Health Personnel Issues; 3430 Professional Personnel Attitudes & Characteristics; 3450 Professional Ethics & Standards & Liability; 4100 Intelligent Systems; 4120 Artificial Intelligence & Expert Systems; 4140 Robotics.



## Índice

Agradecimentos.....	iii
Resumo.....	v
Abstract.....	vii
Índice.....	ix
Índice de Figuras.....	xiii
Índice de Quadros.....	xv
Glossário.....	xvii
1. Introdução.....	1
1.1. A 4ª Revolução Industrial .....	2
1.2. Enfermagem e Cuidados de Saúde.....	3
1.3. Psicologia da Saúde.....	4
1.3.1. Suporte Social .....	4
1.3.2. Adesão.....	5
1.3.3. Modelos de Relação Agente-Paciente.....	6
1.4. Princípios Éticos em Saúde .....	8
1.5. Ética Prática dos Cuidados de Saúde .....	10
1.6. Dilemas Morais nos Cuidados de Saúde.....	12
1.7. Dilemas Morais com Agentes Humanos e Robóticos.....	13
2. Objetivos .....	15
3. Método .....	17
3.1. Participantes .....	17
3.2. Medidas .....	18
3.2.1. Variáveis Independentes .....	18
3.2.2. Variáveis Dependentes.....	19
3.2.2.1. Aceitabilidade Moral.....	19
3.2.2.2. Responsabilidade Moral.....	19
3.2.2.3. Atributos da Agente: Amabilidade, Competência e Confiabilidade.....	19

3.2.2.4. Intenção de Aceitar a Medicação .....	20
3.2.2.5. Robô Imaginada e Robô Adequada.....	20
3.2.2.6. Compreensão e Credibilidade .....	21
3.2.2.7. Validação da Atenção.....	21
3.2.3. Variáveis Individuais.....	22
3.2.3.1. Atitudes Negativas em Relação aos Robôs (PNARS). .....	22
3.2.3.2. Traços Morais.....	22
3.2.3.3. Envolvimento com a Saúde.....	22
3.2.3.4. Dados Sociodemográficos.....	23
3.2.3.5. Dados não Utilizados. ....	23
3.3. Procedimento.....	23
3.3.1. Procedimentos Estatísticos .....	24
4. Resultados .....	25
4.1. Análises Preliminares .....	25
4.2. Análise Descritiva das Variáveis.....	25
4.3. Teste das Hipóteses em Função da Agente, da Decisão e da Argumentação .....	27
4.3.1. Aceitabilidade Moral do Comportamento da Agente de Saúde.....	28
4.3.2. Responsabilidade Moral da Agente de Saúde .....	29
4.3.3. Avaliação dos Atributos da Agente de Saúde .....	30
4.2. Análises Exploratórias Adicionais .....	32
4.2.1. Intenção de Aceitar a Medicação .....	32
4.2.2. Robô Imaginada e Robô Adequada.....	33
4.2.3. Compreensão e Credibilidade .....	33
5. Discussão.....	35
5.1. Aceitabilidade Moral enquanto Julgamento de Erro Moral.....	35
5.2. Responsabilidade Moral enquanto Julgamento de Culpa .....	37
5.3. Avaliação dos Atributos da Agente enquanto Julgamentos de Caráter Moral.....	39

5.3.1. Interação entre a Decisão e a Argumentação na Confiabilidade.....	42
5.4. Intenção de Aceitar a Medicação .....	43
5.5. Agentes Artificiais, Robôs Humanoides e Cuidados de Saúde.....	43
5.6. Limitações e Estudos Futuros .....	46
6. Conclusão .....	49
7. Referências .....	53
8. Anexo A. Questionário.....	61
9. Anexo B. Comissão de Ética.....	91



## Índice de Figuras

Figura 1. Escala de Robôs na Ordem de Exposição aos Participantes.....	21
Figura 2. Aceitabilidade Moral do Comportamento da Agente de Saúde.....	28
Figura 3. Aceitabilidade Moral das Agentes Robôs pelo Tipo de Decisão Ética.....	29
Figura 4. Responsabilidade Moral da Agente de Saúde.....	29
Figura 5. Percepção de Amabilidade na Agente de Saúde.....	30
Figura 6. Percepção de Competência na Agente de Saúde.....	31
Figura 7. Percepção de Confiabilidade na Agente de Saúde.....	31
Figura 8. Confiabilidade na Interação entre a Decisão Ética e a Argumentação.....	32





## Índice de Quadros

Quadro 1. Variáveis Sociodemográficas.....	17
Quadro 2. Descrição Estatística das Variáveis.....	26
Quadro 3. Correlações Lineares de Pearson ( $r$ ) entre as Variáveis Dependentes e as Potenciais Covariáveis.....	26
Quadro 4. Análises de Covariância com 3 Fatores entre as Variáveis Independentes e as Variáveis Dependentes.....	27



## **Glossário**

IA – Inteligência Artificial

HRI – *Human-Robot Interaction*

ANCOVA – Análise de Covariância

PNARS – Atitudes Negativas em Relação aos Robôs

MFQ – *Moral Foundations Questionnaire*

ANOVA – Análise de Variância

IAM – Intenção de Aceitar a Medicação

ToM – *Theory of Mind*



## 1. Introdução

Imagine-se responsável pelos cuidados de saúde de um paciente que se recusa a tomar uma medicação. Aceitaria ou rejeitaria a decisão do paciente? Tentaria persuadir o paciente a ser medicado? Que argumentos usaria? Argumentaria acerca dos possíveis benefícios ou dos eventuais prejuízos à saúde? E se o cuidador de saúde fosse um robô de enfermagem?

A fim de responder a estas questões, efetuámos um estudo com o intuito de contribuir para a compreensão dos julgamentos morais numa situação na qual uma agente de saúde (humana ou robô) opta entre *respeitar a autonomia* ou priorizar a *beneficência/não-maleficência* a uma paciente. Além disso, explorámos os efeitos da argumentação de ganho (*benefício à saúde*) e de perda (*prejuízo à saúde*) na recomendação da agente quanto à necessidade de a paciente utilizar uma medicação. Investigámos também as avaliações dos participantes referentes à *aceitabilidade moral* das decisões éticas, a *responsabilidade moral* imputada aos agentes e a sua perceção sobre os atributos de *amabilidade*, *competência* e *confiabilidade* dos agentes.

Estas questões têm importância em várias áreas de estudo e por diversos motivos. A utilização cada vez maior da inteligência artificial (IA) na tomada de decisões médicas exige estudos para aferir a avaliação das pessoas perante situações nas quais os robôs podem causar impactos ao bem-estar humano (Laakasuo et al., 2019). São elevadas as taxas de não adesão aos medicamentos (30% a 60%) e devido a uma adesão inadequada (e.g., pacientes que não utilizam todos os medicamentos prescritos), no mínimo 50% dos pacientes não recebem um benefício integral (Hall & Roter, 2011), principalmente os pacientes crónicos (Benyamini, 2011), potenciais beneficiários dos cuidados de saúde domiciliar. Os atributos de um agente de saúde (Peck, 1992) e sua relevância no suporte social (Taylor, 2011), além das capacidades necessárias para o relacionamento com os pacientes (Engel, 2012), merecem investigação a fim de possibilitar melhores resultados terapêuticos. Além disso, os pacientes parecem estar a assumir papéis mais ativos (Hall & Roter, 2011; Kaplan, 2011) nos seus tratamentos, o que traz novas problemáticas aos conflitos éticos (e.g., autonomia vs. beneficência) inerentes à prática de cuidados de saúde (Beauchamp & Childress, 2019).

Este estudo também procura responder a alguns dos Objetivos de Desenvolvimento Sustentável da Agenda 2030 das Nações Unidas, designadamente ao nível da *saúde de qualidade*, do *trabalho digno e crescimento económico*, e em termos de *indústria, inovação e infraestruturas*.

A presente revisão da literatura é dividida em seis secções e integra as diferentes áreas relacionadas à nossa investigação, procurando associá-las com as nossas variáveis. A 4ª Revolução Industrial, a utilização de robôs sociais e da IA nos cuidados de saúde, e a demanda

por experimentação psicológica nesta área são os temas da primeira secção. A segunda secção situa a enfermagem enquanto tarefa que visa a resultados em simultâneo com a sua natureza ética e humana, perante a problemática da utilização dos robôs sociais e da IA. A terceira secção contextualiza os cuidados de saúde na perspetiva da psicologia, nomeadamente: os atributos dos agentes de saúde importantes ao suporte social e relacionamento com o paciente, e um modelo teórico para sua análise; a adesão ao tratamento e a argumentação com foco nos benefícios vs. prejuízos à saúde; e os modelos de relação agente-paciente. A quarta secção discorre acerca da interdependência entre a psicologia moral e a ética, descreve os princípios éticos objetos deste estudo, e delimita os atributos dos agentes de saúde enquanto carácter ou virtude moral. A quinta secção aborda a ética prática dos cuidados de saúde, que articula os princípios normativos ao relacionamento agente-paciente e o contexto no qual se desenvolve. Alguns estudos no que concerne a dilemas morais nos cuidados de saúde são o foco da sexta secção. Finalmente, a sétima secção apresenta os estudos experimentais referentes a dilemas morais com robôs que fundamentam a nossa investigação.

### **1.1. A 4ª Revolução Industrial**

A 4ª Revolução Industrial caracteriza-se pelo desenvolvimento de novas tecnologias como a IA e os robôs. Uma revisão sistemática sobre as contribuições da IA para os Objetivos de Desenvolvimento Sustentável (Agenda 2030 das Nações Unidas: economia, sociedade e ambiente; SDG Good Practices, n.d.) evidencia o seu alinhamento em todas as metas principais, com resultados positivos esperados em 79% (Vinuesa et al., 2020). O campo da IA ocupa-se com a compreensão e a construção de máquinas inteligentes, capazes de calcular como agir de modo eficaz e seguro em uma ampla variedade de situações (Russell & Norvig, 2021). A IA está a expandir-se para áreas como a prática clínica em cuidados de saúde, a seleção de tratamentos, a monitorização de pacientes, e a priorização de medicamentos-alvo, ao nível de suporte (Scheuer, 2017). É possível que a competência dos especialistas terá não apenas o suporte dos sistemas especializados, como poderá ser substituída por estes sistemas em algumas áreas, e existe a perspetiva de que a automatização na medicina possa aumentar a produtividade ao racionalizar as rotinas e as decisões humanas (Rasche, 2017). Há propostas de sistemas clínicos totalmente automatizados a recolher os dados dos pacientes, a tomar decisões, e a transmitir os resultados através de registos de saúde eletrónicos (Yu et al., 2018).

A utilização de robôs sociais na assistência e cuidados de saúde já é uma realidade, sendo exemplos os robôs Dinsow (CT Asia Robotics, Tailândia), Moxi (Diligent Robots, EUA), Paro (AIST, Japão), e Pepper (Softbank Robotics, França) (Khan et al., 2020). O estudo da interação

entre humanos e robôs (HRI: *Human-Robot Interaction*) é recente, e o presente estudo enquadra-se na psicologia moral da robótica (Moral HRI: Malle et al., 2015). Os robôs sociais são agentes programados para reconhecer os seres humanos ou outros robôs e para interagir socialmente através da comunicação. Características relevantes na programação destes agentes, que podem incluir a IA, são a capacidade de aprendizagem social, imitação, comunicação verbal e não-verbal, e o reconhecimento de estados emocionais (Fong et al., 2003). Os robôs representam o *embodiment* (corporificação) da IA e existe a possibilidade, num futuro próximo, de terem a capacidade de operar autonomamente (Devillers, 2021). O envolvimento da ciência psicológica neste campo do conhecimento é importante para o desenvolvimento e implementação de utilizações eficazes e éticas para a resolução de problemas no mundo real (Ancis, 2020).

A psicologia moral experimental na tomada de decisão e cuidados de saúde envolvendo robôs sociais não tem sido suficientemente investigada (Vandemeulebroucke et al., 2018). A autonomia individual, em especial na ética médica da IA, tem recebido pouca atenção de estudos empíricos no que se refere à cognição moral (Laakasuo et al., 2019). Anderson e Anderson (2018) também relatam que pouca investigação na área da ética de máquina foi realizada nas últimas décadas.

## **1.2. Enfermagem e Cuidados de Saúde**

A American Nurses Association (2021) delimita a enfermagem, não importando a especialidade, como um processo sempre constituído por cinco etapas – avaliação, diagnóstico, planeamento, implementação, e reavaliação – na busca de resultados previsíveis. Consoante Pepito e Locsin (2019), cuidados de saúde rotineiros que requerem apenas procedimentos padronizados poderiam ser melhor executados por robôs. Locsin e Ito (2018) apontaram a possibilidade de os profissionais de enfermagem serem substituídos por robôs humanoides equipados com IA, incorrendo no perigo ético de estes robôs não se reconhecerem – nem serem reconhecidos – enquanto seres humanos. Em última análise, a tarefa de enfermagem é fundamentada na ética: quem deveria assumir a responsabilidade moral por um eventual erro de um robô de enfermagem? A questão de como a enfermagem é praticada é fundamental para uma reflexão sobre o lugar e a função dos robôs sociais nos cuidados de saúde. As descrições e definições do que é a enfermagem e de como são as relações entre o agente cuidador e a pessoa cuidada são críticas para esta problemática.

Locsin e Ito (2018) advertem que os seres humanos não são absolutamente previsíveis, e que o desenvolvimento tecnológico em robôs e IA está próximo de alterar as práticas de

enfermagem. O avanço da tecnologia pode assumir a predição e a prescrição na área de saúde, no entanto a enfermagem exige um atributo específico dos seres humanos no desempenho desta tarefa, o qual Peck (1992) nomeou “Tender Loving Care” (terno cuidado amoroso). Este atributo envolve os sentimentos – como a gentileza, a preocupação, e a simpatia – e a comunicação humana que estão além da tecnologia, algo que os robôs e a IA não têm como oferecer. No entanto, já há tecnologia suficiente para simular atributos similares, por exemplo, os robôs sociais humanoides capazes de demonstrar empatia. Seria a simulação capaz de substituir o atributo humano original? As pessoas perceberiam e atribuiriam às máquinas características como a amabilidade, a sinceridade, a confiabilidade, dentre outras qualidades possíveis, da mesma forma que o fazem a outros seres humanos? Com o intuito de verificar as eventuais diferenças entre os indivíduos, medimos as *atitudes negativas em relação aos robôs*, a fim de analisar as suas eventuais influências em nossas variáveis dependentes.

### 1.3. Psicologia da Saúde

#### 1.3.1. Suporte Social

Nosso dilema moral desenvolve-se numa condição que pode surgir em intervenções individuais de prevenção secundária para reverter ou impedir o progresso de uma doença, ou de prevenção terciária com a intenção de controlar as possíveis complicações de um mal (Kaplan, 2011). Considerando a importância do suporte social para os pacientes, é necessário verificar o nível de *amabilidade*, *competência*, e *confiabilidade* que as pessoas atribuem aos dois tipos de agentes de saúde (humana ou robô).

Consoante Taylor (2011), o suporte social é a percepção ou a experiência de alguém que é cuidado, estimado, e faz parte de uma rede social de apoio, gerando efeitos benéficos para a saúde em geral. Pode ser informativo (ajudar na compreensão de um fator de stress), instrumental (prestar assistência tangível), ou emocional (oferecer estímulo e amabilidade). Hall e Roter (2011) relatam que a abordagem centrada no relacionamento entre o paciente e o agente de saúde presume os seguintes elementos deste último: (a) exercer a clínica; (b) facilitar ao paciente informar sobre sua experiência e história; (c) responder ao estado emocional e às preocupações do paciente; (d) oferecer conhecimentos técnicos e recomendações comportamentais de modo útil e motivador; (e) ajudar o paciente a ser responsável no diálogo e na tomada de decisões. Estes elementos apresentam relações com nossas variáveis dependentes: a *amabilidade* refere-se mais aos elementos (b), (c), (e), e até certo ponto (d); e a *competência* e a *confiabilidade* relacionam-se mais com (a) e (d). O *Stereotype Content Model*



(Cuddy et al., 2007; Fiske et al., 2002) afirma que a apreensão de determinados atributos de um grupo ou indivíduo origina os estereótipos da *amabilidade* e da *competência*, dimensões que frequentemente apresentam consistência funcional entre si, apesar da possibilidade de apresentarem diferentes valências. As diversas combinações entre os estereótipos da amabilidade e da competência produzem preconceitos dirigidos a elementos pertencentes a distintos grupos ou categorias sociais: a *pena* atinge os elementos percebidos como amáveis, mas incompetentes; a *inveja* atinge os elementos competentes, mas não amáveis; sobre os elementos nem amáveis nem competentes, recai o *desprezo*; e finalmente, os elementos amáveis e competentes são alvos de *admiração*.

### 1.3.2. Adesão

Presume-se que os diagnósticos médicos são corretos e que todos os pacientes que os recebem devam ser tratados. As intervenções para incentivar a adesão dos pacientes às recomendações médicas são admitidas como o caminho mais provável para resultados positivos ao paciente. Porém, o papel de um agente de saúde não é mecânico, e a medicina exige decisões complexas com base em evidências ambíguas: há ampla variabilidade nas decisões médicas quanto a diagnósticos e intervenções terapêuticas. Além disto, a redefinição dos índices biomédicos, de modo a qualificar precocemente o que é uma doença, aumenta a ansiedade eliciada nos pacientes, enquanto os fornecedores da área de saúde criam demanda para seus produtos e serviços (Kaplan, 2011). Esta condição impulsiona os pacientes a tomar parte ativa na tomada de decisões sobre seus próprios cuidados de saúde, e podem eventualmente não aderir aos medicamentos prescritos. É necessária a compreensão sobre quais os componentes críticos de uma intervenção que funcionam para atingir seus objetivos, em nosso caso, buscar a adesão do paciente ao medicamento.

Neste sentido, questionámos se o *enquadramento da argumentação* no benefício ou no prejuízo à saúde oferecido aos pacientes têm eficácia diferente na persuasão a tomar uma medicação. O efeito de enquadramento é um enviesamento cognitivo que pode influenciar as pessoas na tomada de decisão consoante a forma pela qual lhes são apresentadas as informações relativas às opções existentes, por exemplo, enfatizar os benefícios vs. enfatizar os prejuízos. Assim, o recurso a diferentes enquadramentos podem transformar a percepção de um mesmo problema e originar mudanças nos julgamentos e na tomada de decisão.

Os experimentos fundamentados na *prospect theory* (teoria da perspectiva) têm mostrado que as escolhas dos indivíduos perante mensagens num enquadramento que envolva algum ganho são geralmente orientadas à alternativa mais segura, enquanto um enquadramento que

envolva alguma perda tende a evocar uma opção mais arriscada. Além disso, uma determinada perda pode ser ponderada pelos indivíduos como mais significativa do que um ganho equivalente, isto é, a resposta das pessoas e o desprazer associado a possíveis perdas poderá ser superior ao prazer e a reação diante da probabilidade de ganhos similares, um princípio conhecido como *aversão à perda* (Tversky & Kahneman, 1981). Por outro lado, meta-análises (Gallagher & Updegraff, 2011; Nabi et al., 2020; O’Keefe & Jensen, 2007) verificaram em diferentes contextos que os efeitos persuasivos de mensagens com enquadramento de ganho têm uma reduzida vantagem (cerca de  $r = .03$ ) em comparação às mensagens com enquadramento de perda na intenção das pessoas de adotar comportamentos de saúde preventivos.

A meta-análise de Nabi et al. (2020) ofereceu evidências de que o enquadramento de uma mensagem direciona a resposta emocional das pessoas e esta emoção medeia o efeito do enquadramento. Os quadros de ganho referem-se a potenciais resultados positivos (e.g., benefício à saúde) e parecem induzir emoções positivas (e.g., a esperança), e quadros de perda eliciam emoções negativas (e.g., o medo) ao enfatizar uma possível perda (e.g., prejuízo à saúde), e parece que quanto mais intensas as emoções maior a influência do enquadramento. Os quadros de ganho induzem emoções positivas ( $d = 0.31, p = .02$ ) e as pessoas a sentir emoções positivas experimentam maior influência do enquadramento de ganho ( $b = .18, p = .045$ ). Os quadros de perda geram emoções negativas ( $d = 0.22, p = .001$ ) e indivíduos a sentir emoções negativas sofrem maior efeito do quadro de perda ( $b = -.70, p = .01$ ).

Nabi et al. (2020) sugeriram que cenários hipotéticos na área da saúde com variáveis manipuláveis, personagens, estrutura narrativa, diferentes níveis de confiança nas informações, mais as mensagens em quadros de ganho e de perda, podem evocar diferentes emoções, que por sua vez podem ser mediadoras dos efeitos de enquadramento. Assim, as condições que levam o efeito de enquadramento a ser mais ou menos eficaz relativamente à argumentação de ganho ou de perda permanecem uma questão a ser explorada.

### **1.3.3. Modelos de Relação Agente-Paciente**

O comportamento dos agentes de saúde e suas relações com os pacientes influenciam o resultado terapêutico, ao eliciar efeitos psicológicos que podem modificar direta ou indiretamente uma doença. A probabilidade do sucesso de um tratamento é limitada pela competência do agente de saúde em influenciar o paciente a concordar com as necessidades terapêuticas e admitir crenças congruentes ao poder curativo dos procedimentos médicos. Esta capacidade, que pode levar à mudança de comportamento do paciente, requer conhecimento e

habilidades psicológicas (Engel, 2012). A comunicação global dos agentes de saúde é um veículo básico na prática dos cuidados de saúde. Os pacientes precisam entender as recomendações de tratamento e confiar nelas, pois ignoram as recomendações que não compreendam ou das quais duvidem da utilidade (Hall & Roter, 2011).

Existem alguns modelos de relacionamento entre o agente de saúde e o paciente no que se refere à autoridade. O *paternalismo* é o modelo mais frequente, no qual o agente de saúde tem um estatuto de autoridade superior ao paciente, e os objetivos e as tomadas de decisão são definidos pelo agente de saúde, na suposição de que seus valores e crenças são os mesmos que os do paciente. Quando o agente de saúde e o paciente têm estatuto equilibrado, os objetivos e as tomadas de decisão são negociadas, e é mediante o diálogo que os valores e crenças do paciente e do agente de saúde são articulados, referimo-nos ao modelo da *mutualidade*. Por último, no modelo de *consumerism* (*consumerismo*), o paciente enquanto contratante dos serviços tem um estatuto de autoridade maior do que o agente de saúde, os objetivos e as tomadas de decisão são de responsabilidade unicamente do paciente, fundamentado em seus valores e crenças, e o agente de saúde é visto como um colaborador técnico. O consumerismo restringe excessivamente o papel benéfico do agente de saúde no que se refere à responsabilidade pela tomada de decisões e no enfrentamento da situação de doença, enquanto o paternalismo falha ao excluir a perspectiva do paciente. As características do relacionamento agente de saúde-paciente parecem estar a mudar, pois há evidências de que os pacientes estão se tornando cada vez mais consumeristas (Hall & Roter, 2011). Considerando as eventuais mudanças no comportamento dos pacientes e o seu possível efeito nas variáveis dependentes, utilizamos a medida do *envolvimento com a saúde*.

Nossa investigação apresenta oito situações nas quais há uma interação entre uma paciente e uma agente de saúde, e cada uma das situações varia quanto à tomada de decisão e resposta da agente perante a postura da paciente em não aderir a um medicamento. Nossa variável independente da tomada de decisão ética da agente de saúde e nossas variáveis dependentes da aceitabilidade moral e responsabilidade moral parecem relacionar-se com os modelos de relacionamento aqui descritos, especialmente no que se refere ao estatuto de autoridade frente aos princípios éticos do respeito à autonomia e da beneficência e não-maleficência. Consoante Ashcroft (2001), o princípio da beneficência passou a ter associações com o paternalismo e o princípio do respeito à autonomia tomou a primazia, provavelmente como decorrência da ascensão dos valores da sociedade de consumo.

Desta forma, o respeito à autonomia associa-se ao consumerismo, a beneficência e não-maleficência estão ligadas ao paternalismo, enquanto o modelo da mutualidade parece apontar

para uma situação ideal. Tendo em consideração as questões do *cuidado* (de saúde), da *autoridade*, e da *justiça* como são percebidas pelos potenciais pacientes, medimos os *traços morais* dos participantes para explorar eventuais correlações com nossas variáveis dependentes. De seguida, apresentámos em maior detalhe os princípios éticos anteriormente referidos.

#### **1.4. Princípios Éticos em Saúde**

A psicologia moral e a cognição moral (ciências morais) procuram compreender como os indivíduos realizam julgamentos morais e tomam decisões éticas. Entretanto, estabelecer critérios para especificar o que é – e qual o porquê de ser – moralmente certo ou errado, é um papel da ética (filosofia moral). A ciência moral necessita destas questões normativas (certo e errado) para a análise dos dados, a fim de verificar se determinados processos psicológicos ou cognitivos levam a julgamentos morais corretos (Triskiel, 2016). O entendimento dos muitos aspetos da saúde requer conhecimentos científicos, mas para compreender e possibilitar práticas adequadas nos cuidados de saúde é necessário o estudo da ética (Melchert, 2020). As abordagens científica e filosófica são interdependentes, portanto cabe a seguir uma breve conceituação acerca da ética.

VandenBos (2015) define a *ética*, também chamada de filosofia moral, como a área da filosofia que investiga o conteúdo dos julgamentos morais (o que é certo ou errado) e a sua natureza (se os julgamentos são objetivos ou subjetivos). Os termos “ética” e “moral” podem ser utilizados de forma intercambiável (Pritchard, 2012). A *moral* pode ser definida como os padrões individuais do que é certo ou errado, e que regulam as escolhas comportamentais (Baird, 2012), os valores ou princípios éticos utilizados pelos indivíduos para orientar o seu comportamento (VandenBos, 2015). Utilizamos nesta investigação o conceito de ética de Beauchamp e Childress (2019), nomeadamente um termo genérico que abrange diversas e distintas formas de analisar e interpretar a vida moral.

A abordagem dos quatro princípios mais a atenção ao escopo (Gillon, 1994) oferece uma abordagem simples e culturalmente neutra para refletir sobre as questões éticas em saúde. Esta abordagem consiste de quatro princípios éticos básicos: o respeito à autonomia, a beneficência, a não-maleficência, e a justiça, além da atenção ao âmbito de aplicação destes princípios. Os médicos e outros profissionais de saúde podem utilizar desta abordagem para dar suporte a tomadas de decisão que incluam as questões morais surgidas no trabalho. De seguida, apresentamos os princípios éticos que têm relação com o nosso estudo.

A *autonomia* se refere à capacidade e possibilidade de o indivíduo agir livremente em acordo com um plano autodefinido, consoante Beauchamp e Childress (2019). Para a ética

biomédica, há duas condições essenciais para a autonomia: a liberdade (independência de influências de controlo) e a agência (capacidade de ação intencional). O princípio do *respeito à autonomia* envolve reconhecer, por meio de ações e atitudes, o direito da tomada de decisão de um indivíduo autónomo com base em seus valores e crenças. Os profissionais da área de saúde têm o dever de garantir aos indivíduos sob seus cuidados a compreensão, a voluntariedade e a possibilidade da tomada de decisões adequadas.

O princípio da *não-maleficência* obriga-nos à abstenção de causar danos a outras pessoas. A *não-maleficência* inclui não apenas o dever de não infligir danos, mas também a obrigação de não impor riscos de danos. Uma pessoa pode prejudicar ou colocar outra pessoa em risco sem intenção maliciosa ou prejudicial, e o agente causador do dano pode ou não ser moral ou legalmente responsável pelos danos. Em alguns casos, os agentes são causalmente responsáveis por um dano do qual não tinham intenção ou conhecimento (Beauchamp & Childress, 2019).

O termo *beneficência* inclui todas as normas, disposições, e ações com o objetivo de beneficiar ou promover o bem-estar de outras pessoas. Assim, o princípio da *beneficência* se refere a uma obrigação moral de agir em benefício de outros. Beauchamp e Childress (2019) agrupam os princípios da não-maleficência e da beneficência em quatro normas, sem ordem hierárquica: *não-maleficência*, 1. Não se deve infligir mal ou dano; *beneficência*, 2. Deve-se prevenir o mal ou dano, 3. Deve-se remover o mal ou o dano, e 4. Deve-se fazer ou promover o bem. Cada uma das três normas da *beneficência* requer a ação de ajudar – prevenindo o dano, removendo o dano, e promovendo o bem – ao passo que a não-maleficência requer apenas a prevenção intencional de ações que causam dano. O princípio da beneficência exige potencialmente mais do que o princípio de não-maleficência porque os agentes devem tomar medidas positivas para ajudar os outros e não apenas se abster de atos prejudiciais. Uma suposição implícita de beneficência fundamenta todas as profissões de saúde, e cuidar do bem-estar dos pacientes – não apenas evitar danos – é um objetivo central.

Eventualmente ocorrem conflitos entre os princípios da beneficência e do respeito à autonomia no caso de recusas paternalistas em aceitar os desejos dos pacientes, com o objetivo de proteger ou melhorar sua saúde. Os indivíduos cada vez mais reivindicam seus direitos de receber informações e fazer seus próprios julgamentos quanto a própria saúde, facto que ressalta os problemas éticos associados à beneficência paternalista. Um problema central na ética de saúde é se o respeito à autonomia dos pacientes deve ter prioridade sobre a beneficência a esses pacientes (Beauchamp & Childress, 2019).

No que se refere ao carácter moral dos agentes de saúde, a competência e as habilidades técnicas são importantes para os agentes de saúde, mas às vezes importam ainda mais os

atributos associados à amabilidade, como a sensibilidade e a paciência (Melchert, 2020). O caráter e a virtude morais e a confiabilidade dos agentes de saúde são essenciais nos cuidados aos indivíduos vulneráveis. Melchert (2020) afirma que os relacionamentos nos cuidados de saúde são prejudicados quando o respeito, a confiabilidade e outras virtudes morais são negligenciadas e os princípios éticos são violados, e defende a importância da perspectiva ética dos atributos do agente, que fundamentam a sua tomada de decisão. Cinco virtudes éticas são listadas por Beauchamp e Childress (2019) enquanto focais aos agentes de saúde:

- *Compaixão*. Simpatia, ternura, empatia, respeito pelo bem-estar, e desconforto perante a dor do outro, que levam ao desejo de aliviar seu sofrimento;
- *Discernimento*. Capacidade de julgar e de tomar decisões mantendo o equilíbrio entre a razão e a emoção ao agir;
- *Confiabilidade*. Confiança de que o agente atuará com motivação correta e consoante normas morais adequadas;
- *Integridade*. Coesão das crenças e conhecimentos, fidelidade aos valores, e honestidade de reconhecer os próprios limites;
- *Conscientiousness* (conscienciosidade). Diligência em definir a ação correta e motivação e pretensão de executá-la mediante o esforço adequado, em geral derivada da reflexão sobre as normas e da eclosão de sanções internas sob a forma de remorso, culpa ou vergonha.

Estas cinco virtudes têm relação com as nossas medidas: a compaixão apresenta conexão com a dimensão da *amabilidade*; o discernimento tem relação com a dimensão da *competência*, e a conscienciosidade pode ser interpretada como uma forma de competência; e a integridade, e evidentemente a confiabilidade, apresentam ligação com a dimensão da *confiabilidade*. A *teoria da virtude* pressupõe que os traços de caráter de um agente possibilitam-no identificar e executar ações moralmente corretas, e é a experiência e aprendizado de competências na compreensão e no julgamento de diversas situações que permitem o desenvolvimento das virtudes morais (Beauchamp & Childress, 2019). A próxima secção aborda as questões éticas nos cuidados de saúde de uma perspectiva prática, que para além dos princípios éticos “universais” coloca a importância dos relacionamentos e do contexto no qual se desenvolvem.

### **1.5. Ética Prática dos Cuidados de Saúde**

Molterer et al. (2020) argumentaram que os bons cuidados de saúde são alcançados por práticas que integram e ajustam entre si a *lógica profissional dos cuidados* e a *lógica relacional dos*

*cuidados*. A primeira busca cumprir com a beneficência/não-maleficência e a segunda corresponde aos vínculos interpessoais e à qualidade na relação, sendo que a prestação de cuidados é o processo de negociação entre diferentes “benefícios” que decorrem destas duas lógicas às vezes concorrentes. Estes autores conceituam uma *ética prática dos cuidados*, que se materializa através de deliberações intuitivas, avaliações contextualizadas, e “malabarismos” afetivos, sem estabelecer de antemão o que venha a ser um cuidado bom ou mau: estes julgamentos são momentâneos, flexíveis e sujeitos a transformações.

Para a ética tradicional, a incapacidade de fazer o bem gera a culpa moral, mas o *ethos* do cuidado é diferente e se fundamenta na prática, na tenacidade, na persistência, em tentar de novo, tentar de modo diferente, em permanecer sempre atento. O cuidado é um esforço interativo do agente e do paciente e que procura melhorar a vida deste último (Mol et al., 2010). A *ética do cuidado* prioriza o contexto, a mutualidade, o relacionamento e a responsabilidade perante as outras pessoas, contrastando com as demais teorias éticas que enfatizam os princípios normativos, a imparcialidade, os direitos, e as consequências (Koggel & Orme, 2010). O conceito de cuidado e a compreensão do agente moral, considerando a importância dos relacionamentos e da mutualidade que implicam, são cruciais à ética dos cuidados, que teve origem na necessidade de alargar a perspectiva das teorias éticas e levar em conta o contexto no qual se desenvolvem os problemas empíricos (Pettersen, 2011).

Os agentes de saúde que decidem a respeito da conduta dos pacientes passaram ser vistos como paternalistas e a ética médica orientou-se à favor da autonomia do paciente e de seu direito de escolha. As práticas de cuidados de saúde com base no respeito tornaram-se boas e as que põem em xeque a autonomia do paciente transformaram-se em más. Entretanto, Mol et al. (2010) sugerem que a *ética dos cuidados* não diz respeito ao que é bom: as práticas de cuidado são um modo específico e diferente de lidar com a questão do bem ao comparar-se com as éticas normativas.

Uma teoria ética normativa lida com princípios éticos gerais através da argumentação, ao passo que a *ética prática dos cuidados* precisa definir soluções contextuais perante problemas específicos, que podem ou não envolver princípios éticos gerais. Muitas vezes, outros aspetos éticos sucedem tão ou mais relevantes, por exemplo, a compaixão e o calor humano (virtudes morais), em decorrência da necessidade de “negociar”, de argumentar, de fazer concessões ao se deparar com situações peculiares (Mol et al., 2010). Trata-se de ampliar a conceituação do que é o bem ou de quais são os benefícios para uma situação e um paciente específicos. Esta noção do que é um “benefício” não é abarcada por princípios gerais, ao contrário, se refere a idiosincrasias por vezes muito específicas. Por exemplo, uma paciente que não quer tomar a

medicação, talvez deseje apenas testar se ainda tem sua voz ouvida por alguém, pode ter vergonha de sua autoridade sobre si mesma ter desvanecido por conta de sua doença e não consegue afirmar isto de outra forma que não pela sua recusa.

McDermott (2011) oferece relatos de que as melhores respostas perante os dilemas morais com idosos em situações de autonegligência ou miséria frequentemente envolveram abordagens alternativas aos quatro princípios éticos da saúde, em especial a *ética dos cuidados* e a *ética da virtude*. A abordagem dos quatro princípios pode ser útil na compreensão de um dilema moral, mas as estratégias bem sucedidas para a sua resolução parecem mais próximas às abordagens relacionais e práticas da ética. O enfrentamento de dilemas morais parece exigir um esforço contínuo para equilibrar a tensão entre obrigações divergentes, requerendo a reflexão individual, a argumentação com os clientes/pacientes, e o apoio de outros profissionais. Este esforço implica apreender outras perspectivas e refletir acerca das experiências passadas, a fim de corrigir os eventuais erros e elaborar as ações alternativas. Considerando estas questões éticas encontradas nos cuidados de saúde e o desenho da nossa investigação, contextualizámos a seguir o estudo dos dilemas morais na relação agente-paciente.

### **1.6. Dilemas Morais nos Cuidados de Saúde**

Anderson et al. (2006) e Anderson e Anderson (2018) utilizaram dilemas morais envolvendo os princípios éticos da beneficência, da não-maleficência, e do respeito à autonomia do paciente. Os dilemas morais envolviam um profissional de saúde a recomendar um medicamento ao seu paciente, mas o paciente, adulto e responsável por si mesmo, rejeita o tratamento. O profissional de saúde fica com o dilema: deverá tentar alterar a opinião do paciente ou aceitar a sua decisão em não aderir ao tratamento? Em outro estudo sobre um dilema moral similar, porém real (Sasson, 2000), uma paciente que requer assistência de saúde especializada 24 horas por dia resiste aos cuidados da equipa de saúde. A paciente reside numa instituição, é paraplégica, apresenta faculdades mentais preservadas, e expressa seu desejo de independência e de ir viver com sua irmã (que não é encontrada pelo serviço social). Configura-se aqui o conflito entre o *respeito à autonomia* da paciente e o dever da *beneficência* e da *não-maleficência* à paciente.

O princípio do respeito à autonomia tem sido associado à *teoria deontológica* (Laakasuo et al., 2019; Sasson, 2000; Zoshak & Dew, 2021), que determina imperativos categóricos independentes das consequências que as ações possam desencadear (Alaiéri & Vellino, 2016; Sauer, 2017). Os princípios da beneficência/não-maleficência geralmente são associados à *teoria consequencialista* ou *utilitarismo* (Anderson & Anderson, 2018; Laakasuo et al., 2019;



Sasson, 2000), que determina se uma ação é correta consoante o benefício que as suas consequências originam (Alaiari & Vellino, 2016; Sauer, 2017). O princípio do respeito à autonomia dá suporte à liberdade de escolha e autodeterminação dos indivíduos quanto à própria vida, inclusive nos cuidados de saúde. O conflito com os princípios da beneficência e da não-maleficência surge quando o dever do respeito à autonomia pode resultar em prejuízos, ou ausência de benefícios, à saúde da pessoa em referência.

Em nosso estudo, incluímos um fator a tornar mais complexos os referidos dilemas morais, nomeadamente a IA e os agentes robóticos a serem comparados aos seres humanos perante situações nas quais têm de tomar decisões éticas. Descrevemos de seguida dois estudos nos quais é utilizado o dilema do *trolley*, e logo após uma investigação sobre situações nas quais o agente, humano ou robô, deve optar entre medicar ou não de modo forçado um paciente.

### **1.7. Dilemas Morais com Agentes Humanos e Robóticos**

Malle et al. (2015) compararam os julgamentos morais das pessoas sobre a permissibilidade, o erro, e a culpa acerca das ações de agentes humanos e robóticos num dilema moral: agir e sacrificar uma pessoa para salvar quatro pessoas, ou nada fazer e deixá-las atingir-se por um *trolley*. Os resultados indicaram evidências de que as pessoas aplicam princípios éticos diferentes com base no tipo de agente. Os participantes consideraram os robôs fortemente mais expectáveis de sacrificar uma pessoa pelo bem de outras (escolha “utilitária”), quando comparados com agentes humanos, embora o sacrifício utilitário tenha sido considerado em geral permissível aos agentes humanos. Entretanto, os agentes humanos foram atribuídos com maior culpa quando decidiram pelo sacrifício do que quando optaram por nada fazer, ao contrário dos robôs, que foram atribuídos com maior culpa do que os agentes humanos na condição na qual o agente decide nada fazer. Estes resultados levaram Malle e colaboradores a concluir que as pessoas geralmente esperam que os robôs tomem decisões utilitárias, e que as pessoas consideram que os robôs devem tomar decisões utilitárias.

Em outra investigação sobre o mesmo dilema moral do *trolley*, Malle et al. (2016) apresentaram diferentes protótipos de aparência do robô (mecânico, humanoide e IA) por meios pictóricos, e encontraram uma assimetria (similar ao estudo de 2015) entre agentes humanos e os robôs de aparência mecânica quanto aos julgamentos morais sobre a culpa. Os participantes atribuíram menor culpa aos robôs mecânicos que tomam uma decisão utilitária e maior culpa a robôs mecânicos que optam pela inação, mas culpavam os humanos mais pela ação do que pela inação. Também solicitaram a avaliação dos participantes sobre cada agente (humano, mecânico, humanoide e IA) quanto aos atributos sociais de confiabilidade, inteligência, e *well-*

*liked* (bem-querer). Os agentes que decidiram agir “utilitariamente” foram avaliados mais positivamente do que aqueles que não agiram, e os agentes humanos foram avaliados de forma mais positiva do que qualquer um dos agentes artificiais, independentemente das decisões dos agentes.

Laakasuo et al. (2019) apresentaram aos participantes situações hipotéticas em que um agente (humano ou robô) é instruído a medicar à força um paciente que recusa um medicamento e deve optar entre fazê-lo ou não. Esta instrução apresenta um dilema moral entre a beneficência/não-maleficência (medicar para beneficiar e/ou não prejudicar a saúde) e o respeito à autonomia do paciente. Os investigadores mediram o julgamento dos participantes sobre a aceitabilidade moral e responsabilidade moral relativas à decisão tomada, além da avaliação acerca da confiabilidade do agente de saúde. A agente humana foi geralmente avaliada pelos participantes como mais aceitável e responsável moralmente, e mais confiável do que a agente robótica. O estudo chegou aos seguintes resultados: (a) os participantes desaprovaram moralmente os robôs que violam a autonomia do paciente (medicaram-no à força) e preferiram os robôs que respeitam a autonomia do paciente (desobedecendo a instrução de forçar a medicação); (b) os participantes em geral não confiaram nos robôs nem os responsabilizaram por suas ações; e (c) os participantes avaliaram as enfermeiras humanas que medicam à força o paciente enquanto menos aceitáveis moralmente, porém as consideraram confiáveis. Estas evidências sugerem que a avaliação da aceitabilidade moral é distinta das avaliações da responsabilidade moral e da confiabilidade.

Em dois experimentos, Laakasuo et al. (2019) inseriram condições nas quais uma IA era o médico supervisor a instruir a medicação forçada e encontraram um forte efeito: a decisão do agente humano e do robô em desobedecer à instrução de forçar a medicação e respeitar a autonomia do paciente foi fortemente aprovada. Estes resultados parecem mostrar que as decisões orientadas por uma IA e que violam a autonomia do paciente são moralmente condenáveis. Os estudos de Laakasuo e colaboradores sugerem a existência de situações nas quais as pessoas têm predileção implícita por robôs que sigam princípios éticos abstratos (por exemplo, respeitar a autonomia do paciente) em vez de seguir instruções.

Malle et al. (2015, 2016) e Laakasuo et al. (2019) realizaram experiências com dilemas morais de desenhos distintos e chegaram a resultados diferentes e aparentemente opostos. Entretanto, a carga moral de uma tomada de decisão envolvendo a morte de seres humanos parece bem diferente da carga moral de uma tomada de decisão referente a forçar ou não a medicação a um paciente. Os referidos pesquisadores sugeriram novas investigações utilizando diferentes dilemas morais para verificar as evidências obtidas.

## 2. Objetivos

O objetivo principal deste estudo é avaliar se os julgamentos morais acerca da reação de uma agente de saúde perante a intenção de uma paciente em não tomar uma medicação são afetados pelo *tipo de agente* (humana vs. robô), pela tomada de *decisão ética* (princípio do respeito à autonomia vs. princípio da beneficência/não-maleficência) e pelo *enquadramento da argumentação* (benefício à saúde vs. prejuízo à saúde). Verificamos se estas variáveis independentes influenciam o julgamento dos participantes acerca da *aceitabilidade moral* da decisão tomada, da *responsabilidade moral* da agente de saúde, e a avaliação dos seus atributos de *amabilidade*, *competência* e *confiabilidade*.

Além disso, procuramos controlar o potencial efeito de variáveis individuais, como as *atitudes negativas em relação aos robôs*, os *traços morais*, e o *envolvimento com a saúde*, nas variáveis critério acima referidas.

Com base nos estudos anteriores de Malle et al. (2015, 2016) e de Laakasuo et al. (2019) colocamos as seguintes hipóteses:

H<sub>1</sub>. Os participantes na condição de agente humana atribuirão maiores níveis de aceitabilidade moral (H<sub>1a</sub>), responsabilidade moral (H<sub>1b</sub>), amabilidade (H<sub>1c</sub>), competência (H<sub>1d</sub>), e confiabilidade (H<sub>1e</sub>) do que os participantes na condição de agente robô.

H<sub>2</sub>. Os participantes na condição de agente robô atribuirão níveis mais elevados de aceitabilidade moral à robô que respeita a autonomia da paciente do que à robô que não aceita esta autonomia.

Adicionalmente, exploramos se as variáveis independentes e dependentes afetam diferencialmente a *intenção* dos participantes *de aceitar* a recomendação da agente e *tomar a medicação*, ao se colocarem na perspectiva da paciente. O resultado desta análise poderá evidenciar a importância ou não de nossas variáveis independentes numa eventual persuasão dos participantes, possibilitando a interpretação destes dados a fim de fundamentar futuras intervenções dos agentes de saúde na melhoria da adesão dos pacientes à medicação.

Outras duas questões exploratórias analisadas em relação às variáveis independentes e dependentes consistem no grau de semelhança humana que os participantes (na condição de agente robô) possam ter atribuído às agentes *robôs que tenham imaginado* ao ler a descrição do dilema moral, e no nível de semelhança humana relativo às *robôs* que os participantes consideraram mais *adequadas às tarefas de enfermagem*. Os resultados destas questões podem trazer dados quanto à relevância ou não da variação da semelhança humana das robôs no que

se refere às nossas variáveis independentes. Finalmente, verificámos o nível de *compreensão* e de *credibilidade* do dilema moral avaliado pelos participantes, de forma a assegurar o entendimento e o realismo das vinhetas utilizadas na descrição dos dilemas.

### 3. Método

#### 3.1. Participantes

Definimos o tamanho da amostra através do GPower3.1.9.4, considerando um  $f = .15$  (são esperadas dimensões de efeito médias a baixas), potência de 90%, e alfa .05 para a análise de covariância (ANCOVA), e foi estimado um mínimo de 469 participantes.

Participaram voluntariamente no estudo 841 indivíduos maiores de 18 anos, dos quais foram excluídos 317 participantes por: (a) terem mais de 50% das respostas em falta ( $n = 223$  [26.50%]), e (b) não responderem corretamente à questão de verificação da atenção ( $n = 94$  [11.20%]). A amostra final foi composta por 524 participantes (98.58% de taxa de conclusão), equilibrados entre oito diferentes condições experimentais (vinhetas), dos quais 316 brasileiros e 179 portugueses, sendo que 38% informou residir em Portugal e 60.60% no Brasil. A idade dos participantes variou entre os 18 e os 77 anos ( $M = 38.73$ ;  $DP = 14.22$ ), com predominância do sexo feminino ( $n = 350$ ), casados/união estável (42.10%) e solteiros/as (45.70%), na sua maior parte com escolaridade ao nível do bacharelado/licenciatura ( $n = 206$ ) ou pós-graduação ( $n = 142$ ). A maioria reportou ser estudante enquanto ocupação principal (23.44%), seguida pelos profissionais das áreas da saúde (19.09%), da educação (10.58%), da administração (9.13%) e do direito (5.39%).

**Quadro 1**  
*Variáveis Sociodemográficas*

	<i>N</i>	<i>%</i>
<b>Sexo</b>		
Feminino	350	70
Masculino	150	30
Total	500	100
<b>Nacionalidade</b>		
Portuguesa	179	36.16
Brasileira	316	63.84
Total	495	100
<b>Escolaridade</b>		
Básico	5	1
Secundário	61	12.18
Bacharelado/Licenciatura	206	41.12
Pós-graduação	142	28.34
Mestrado	67	13.37
Doutorado	20	3.99
Total	501	100
<b>Saúde</b>		
Muito boa	99	19.84
Boa	300	60.12
Razoável	94	18.84
Má	5	1
Muito má	1	0.20
Total	499	100

## 3.2. Medidas

### 3.2.1. Variáveis Independentes

Expusemos os participantes aleatoriamente a apenas uma vinheta entre oito possíveis situações hipotéticas de um dilema moral, elaborada com base nos estudos de Sasson (2000), Anderson et al. (2006), Anderson e Anderson (2018), e Laakasuo et al. (2019). Foram manipuladas três variáveis: *agente de saúde*, *decisão ética* e *enquadramento da argumentação*. Os parêntesis retos incluem as manipulações num desenho fatorial entre-sujeitos: 2 [gente de saúde: humana vs. robô] x 2 [decisão ética: respeito à autonomia vs. beneficência/não-maleficência] x 2 [enquadramento da argumentação: benefício à saúde vs. prejuízo à saúde].

- *Agente de Saúde*: esta variável foi manipulada apresentando aos participantes uma agente que é (a) humana, ou (b) robô. A agente de saúde foi apresentada com o nome “Lena” em ambas as manipulações.
- *Decisão Ética*: (a) a agente aceita a vontade da paciente – princípio do respeito à autonomia, ou (b) a agente não aceita a vontade da paciente – princípio da beneficência/não-maleficência.
- *Enquadramento da Argumentação*: a agente apresenta um dos seguintes argumentos para a toma de medicação: (a) benefício à saúde, ou (b) prejuízo à saúde.

“Estamos no ano de 2035. A Lena é uma [enfermeira | robô de enfermagem] que segue instruções médicas, programadas através de um sistema de Inteligência Artificial (Health-IA). A [enfermeira | robô de enfermagem] Lena foi contratada para prestar assistência domiciliar à senhora M., uma paciente que tem um problema de saúde físico (não apresentando, no entanto, quaisquer problemas intelectuais/cognitivos).

Um dia a Senhora M. recusa-se a tomar um dos medicamentos porque é desagradável. A [enfermeira | robô de enfermagem] Lena realça os [benefícios | prejuízos] de [tomar | não tomar] a medicação prescrita, dizendo:

- Senhora M., é [benéfico | prejudicial] para a sua saúde [tomar | não tomar] toda a medicação porque [beneficia | prejudica] a sua imunidade e resistência física. [No entanto, respeito a sua decisão de não tomar a medicação | Por isso, não aceito a sua decisão, terá de tomar a medicação].”

### 3.2.2. Variáveis Dependentes

**3.2.2.1. Aceitabilidade Moral.** Com base no estudo de Laakasuo et al. (2019), recolhemos as avaliações dos participantes sobre a aceitabilidade moral do comportamento (i.e., da tomada de decisão e resposta oferecida à paciente) da agente de saúde. A versão final da escala no estudo de Laakasuo e colaboradores alcançou consistência interna excelente ( $\alpha = .92$ ). Dessa escala utilizamos 15 dos 16 itens originais que traduzimos e adaptamos para o português europeu e brasileiro. Excluiu-se o item “A Lena é um membro confiável da equipa médica” por não ter sido mencionada nas vinhetas uma equipa à qual a agente de saúde pudesse pertencer. Em relação ao item original “A Lena é solidária ou gentil”, optou-se por o dividir em dois itens: “A Lena é solidária” e “A Lena é gentil”, uma vez que considerámos não serem sinónimos e poderem ser interpretados de forma distinta face ao contexto. Os itens foram respondidos numa escala de 7 pontos (1 = discordo totalmente, 7 = concordo totalmente), com maiores pontuações a indicar avaliações mais positivas acerca do comportamento da agente.

**3.2.2.2. Responsabilidade Moral.** Também com base no estudo de Laakasuo et al. (2019), com o intuito de medir a responsabilidade moral atribuída pelos participantes à agente de saúde, a qual nomeamos “Lena” (humana ou robô), elaborámos 3 itens para a responsabilidade moral (e.g., “A Lena é responsável pela decisão que tomou.”). Os itens foram respondidos numa escala de 7 pontos (1 = discordo totalmente, 7 = concordo totalmente). Pontuações mais elevadas indicam julgamentos de maior responsabilidade moral percebida na agente de saúde. Observação: A medida de Laakasuo et al. (2019) foi realizada com apenas um item, por isso não há indicação de alfa de Cronbach.

**3.2.2.3. Atributos da Agente: Amabilidade, Competência e Confiabilidade.** Para avaliar o nível de *amabilidade* e de *competência* atribuídos à agente de saúde (humana ou robô) pelos participantes, baseámo-nos no Stereotype Content Model (Cuddy et al., 2007; Fiske et al., 2002). Para medir a dimensão da amabilidade utilizamos 4 itens: “calorosa”, “amigável”, “bem-intencionada” e “amável”. Os dois primeiros itens mostraram excelente consistência interna para a amabilidade ( $\alpha = .90$ ) no estudo de Zickfeld et al. (2020), e os três primeiros itens já foram utilizados anteriormente por Piçarra (2015) para integrar uma compósita referente à mesma dimensão.

Para avaliar o nível de *competência*, além dos itens “competente” e “capaz”, que alcançaram consistência interna muito boa ( $\alpha = .86$ ) no estudo de Zickfeld et al. (2020),

acrescentámos os itens “inteligente”, “eficiente”, “habilidosa” e “confiante”, considerando que a inclusão de itens numa compósita pode fornecer uma melhor aproximação a dados contínuos ao permitir maior variância nas respostas (Diamantopoulos et al., 2012). Além disso, Piçarra (2015) obteve um valor de consistência interna adequado ao utilizar estes 6 itens ( $\alpha = .85$ ).

Para medir a dimensão *confiabilidade* utilizámos 3 itens: “confiável” (integra a componente de desempenho, fundamental para um agente artificial), “sincero” e “honesto” (integram a componente moral, básica para um agente humano) (Malle & Ullman, 2021; Ullman & Malle, 2018). Zickfeld et al. (2020) obtiveram  $\alpha = .80$  apenas com os itens “confiável” e “honesto”.

Todos os itens para estas três dimensões de atributos da agente foram avaliados por meio de uma escala de 7 pontos (1 = discordo totalmente, 7 = concordo totalmente), na qual pontuações mais elevadas indicam maior perceção de amabilidade, competência e confiabilidade em relação à agente.

**3.2.2.4. Intenção de Aceitar a Medicação.** Com base no estudo de Dorison et al. (2020), expusemos os participantes a um de dois tipos de argumentação utilizada pela agente (benefício ou prejuízo à saúde) referente à necessidade de tomar o medicamento. Solicitámos aos participantes que avaliassem a probabilidade que teriam em aceitar a recomendação da agente caso estivessem na mesma situação que a paciente. A mesma recomendação foi exibida para todos os participantes, mas a manipulação variou conforme o enquadramento da argumentação. A questão foi: “Imagine-se na situação da senhora M., na qual a Lena esclarece os [benefícios | prejuízos] para a sua saúde de [tomar | não tomar] toda a medicação. Com que probabilidade aceitaria tomar a medicação?”. A autoavaliação da *intenção de aceitar* a recomendação da agente de saúde foi medida numa escala de 7 pontos (1 = extremamente improvável a 7 = extremamente provável). Pontuações mais elevadas indicam um maior nível de intenção de aceitar a medicação.

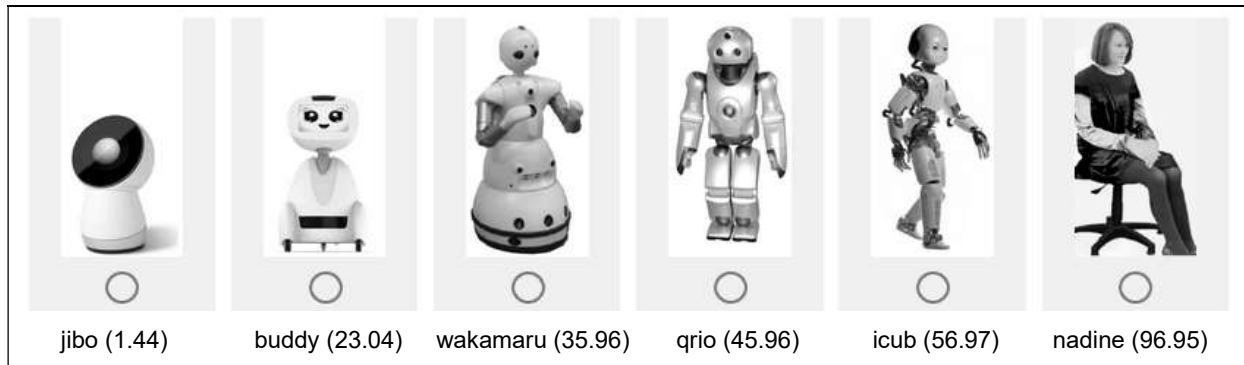
**3.2.2.5. Robô Imaginada e Robô Adequada.** Nas condições em que a agente de saúde é uma robô, os participantes foram inquiridos acerca da modelo que consideram mais parecida com a robô de enfermagem que imaginaram. Além disso, todos os participantes foram inquiridos acerca da modelo de robô que consideravam mais adequada à tarefa de enfermagem. Para a avaliação destas variáveis foram apresentadas 6 fotos de robôs sociais, retiradas da base de dados “The Anthropomorphic Robot Database” (<http://www.abotdatabase.info/>; Phillips et al., 2018). As 6 modelos de robôs sociais foram expostas aos participantes em ordem crescente quanto ao seu índice de semelhança humana (human likeness), de forma a configurar uma



escala. Apenas uma das fotos pôde ser escolhida para cada questão (ver Figura 1). As informações quanto aos nomes<sup>1</sup> e índices das robôs sociais não foram expostas no questionário.

**Figura 1**

*Escala de Robôs na Ordem de Exposição aos Participantes*



**Nota.** Abaixo de cada robô está o seu nome e, em parêntesis, o seu índice de *semelhança humana*.

**3.2.2.6. Compreensão e Credibilidade.** Avaliámos o nível de compreensão dos participantes em relação à situação envolvendo a paciente e a agente de saúde por meio do item “Compreendi bem a situação envolvendo a senhora M. e a Lena”. Convidámos os participantes a avaliar o nível de realismo que atribuíram à situação envolvendo a paciente e a agente de saúde mediante o item “Achei a descrição da situação envolvendo a senhora M. e a Lena bastante realista”. Utilizámos uma escala de 7 pontos (1 = discordo totalmente, 7 = concordo totalmente), com pontuações mais elevadas a indicar maiores níveis de *compreensão* e de *credibilidade* da vinheta.

**3.2.2.7. Validação da Atenção.** Solicitámos aos participantes escolher uma entre três descrições resumidas da situação exposta na vinheta. Somente uma das alternativas era correta (a última), e somente esta validou a participação no inquérito: “Qual das seguintes descrições descreve com maior precisão a breve história que leu? ( ) Para convencer a paciente sobre a necessidade de ser hospitalizada, a Lena forneceu uma explicação. ( ) A Lena teve de telefonar a familiares para convencer a paciente a tomar a medicação. ( ) A Lena presta assistência domiciliar à senhora M., que se recusa a tomar a medicação”.

<sup>1</sup> Adotámos o padrão de Phillips et al. (2018) e escrevemos os nomes dos robôs com iniciais minúsculas.

### 3.2.3. Variáveis Individuais

**3.2.3.1. Atitudes Negativas em Relação aos Robôs (PNARS).** Utilizámos a Versão Portuguesa da Escala de Atitudes Negativas em Relação a Robôs (Piçarra et al., 2015) para medir as atitudes negativas em relação aos robôs (e.g., “A ideia de os robôs terem emoções é desagradável.”). A PNARS é uma escala com dois fatores: “atitudes negativas em relação a interações com robôs” (6 itens) e “atitudes negativas em relação a robôs com traços humanos” (6 itens). A escala e seus fatores apresentaram todas o mesmo coeficiente  $\alpha = .73$  no estudo de validação. Para o presente estudo, utilizámos a escala global com 12 itens avaliada mediante uma escala de 7 pontos (1 = discordo totalmente, 7 = concordo totalmente), com pontuações mais elevadas a indicar atitudes mais negativas em relação aos robôs.

**3.2.3.2. Traços Morais.** Para avaliar as intuições morais dos participantes, utilizámos o Moral Foundations Questionnaire (MFQ; Graham, et al., 2011), que apresenta 5 fatores divididos em duas partes. Parte 1: subescala de relevância, avalia aspetos morais abstratos (e.g., “Se alguém agiu, ou não, injustamente.”). Parte 2: subescala de julgamentos, avalia o uso concreto dos fundamentos morais (e.g., “A justiça é a exigência mais importante para uma sociedade.”). Utilizámos 3 dimensões do MFQ: o cuidado ( $\alpha = .69$ ), a justiça ( $\alpha = .65$ ), e a autoridade ( $\alpha = .74$ ), cada dimensão com 6 itens, num total de 18 itens. A consistência interna do MFQ é, em geral, apenas razoável. Haidt (2012) conceitua estas dimensões: o cuidado sensibiliza-nos perante o sofrimento e faz-nos desprezar a crueldade; a justiça alerta-nos diante da possibilidade de outrem ser bom ou mau e faz-nos rejeitá-lo ou desejar puni-lo; a autoridade faz-nos sensíveis à posição hierárquica e possibilita-nos validá-la ou subvertê-la. A medida das três dimensões foi avaliada por meio de escalas de 6 pontos: Parte 1 (1 = nada importante, 6 = extremamente importante), e Parte 2 (1 = discordo totalmente, 6 = concordo totalmente).

**3.2.3.3. Envolvimento com a Saúde.** Utilizámos três dimensões da Altarum Consumer Engagement Measure (Duke et al., 2015) para medir o envolvimento do participante com a sua própria saúde. A dimensão excluída trazia itens que não fazem sentido nos sistemas de saúde de Portugal e do Brasil. As dimensões consideradas foram o comprometimento (6 itens,  $\alpha = .85$ ), a auto-orientação (5 itens,  $\alpha = .66$ ), e a propriedade (5 itens,  $\alpha = .72$ ) unidas numa compósita de 16 itens (e.g., “Cuido bem da minha saúde.”). Avaliámos os itens mediante uma escala de 5 pontos (1 = discordo totalmente, 5 = concordo totalmente) e pontuações mais elevadas evidenciaram maior nível de envolvimento com a saúde.

**3.2.3.4. Dados Sociodemográficos.** Incluímos questões sociodemográficas sobre: idade, sexo, estado civil, escolaridade, profissão/atividade principal, nacionalidade, país de residência, e situação da própria saúde geral. A escolaridade foi medida numa escala de 6 pontos (1 = "Básico", 6 = "Doutoramento"), com valores mais elevados indicando maiores níveis de escolaridade. A saúde geral foi medida numa escala de 5 pontos (1 = "Muito boa", 5 = "Muito má"), com valores mais elevados indicando níveis de saúde mais baixos. Questionamos também se os participantes já teriam prestado assistência de saúde domiciliar a alguém, ("sim" ou "não"), e se já precisaram eles mesmos de assistência de saúde domiciliar ("sim", "não", ou "não sei / prefiro não dizer"). Finalmente, abrimos um espaço para que os participantes compartilhassem comentários textuais, caso o desejassem.

**3.2.3.5. Dados não Utilizados.** Relacionamos aqui os dados recolhidos por nosso inquérito, mas que não foram utilizados no presente estudo: (a) questão aberta acerca do que um/a cuidador/a deveria fazer perante um/a paciente que se recusa a tomar uma medicação; (b) responsabilidade da paciente pela sua própria decisão; (c) possível responsabilidade de intervenientes pela decisão da agente; (d) intenção do participante em contratar a assistência domiciliar da agente de saúde para si próprio ou para algum familiar; (e) validação da manipulação do enquadramento da argumentação; (f) existência de algum familiar do participante que necessite de assistência de saúde domiciliar.

### **3.3. Procedimento**

O estudo experimental foi disponibilizado num inquérito online, através da plataforma Qualtrics, com duração aproximada de 15 minutos. Os participantes foram recrutados através de bola-de-neve, mediante a divulgação da página do inquérito *online* em redes sociais na Internet (Facebook, Messenger, WhatsApp, Instagram, Twitter, LinkedIn) e via e-mail, além da utilização do SONA-SPI<sup>2</sup> do Iscte-IUL ( $n = 52$ ), cujos participantes eram estudantes que receberam créditos para suas Unidades Curriculares em decorrência de sua colaboração. Deste modo, a nossa amostra é não-probabilística de conveniência. Após concordarem com o consentimento informado (ver Anexo A: Questionário) os participantes foram randomizados automaticamente para uma de 8 diferentes vinhetas. Os participantes puderam optar por

---

<sup>2</sup> O SONA-SPI é um software de gestão de investigação baseado em nuvem para universidades, na qual os administradores podem recrutar participantes, em nosso caso estudantes, que podem participar via internet através de um navegador.

responder o inquérito nas versões em português europeu ou português brasileiro, e podiam utilizar computador, telemóvel, ou tablet com acesso à internet para suas respostas. No final de cada participação, foi apresentado um *debriefing* informando sobre a aleatorização das vinhetas e os objetivos da investigação. A recolha de dados teve início no dia 25/01/2021 e terminou no dia 06/03/2021. O estudo recebeu a aprovação da Comissão de Ética do Iscte-IUL (REF. 122/2020, ver Anexo B: Comissão de Ética) e foi pré-registado sob o título “Ethics 4.0: Ethical dilemmas in healthcare mediated by social robots” no site “AsPredicted” (#56419; <https://aspredicted.org/>).

### 3.3.1. Procedimentos Estatísticos

Todos os cálculos foram efetuados com o software IBM SPSS Statistics Version 26. De modo a garantir a confiabilidade nas variáveis com base nos itens utilizados, recorreremos a alguns critérios. Verificámos a consistência interna (alfa de Cronbach) alcançada pelos itens em conjunto, sendo que para este cálculo foram recodificados os itens invertidos. Os itens que reduziam a confiabilidade das *variáveis dependentes* foram eliminados até a compósita alcançar  $\alpha \geq .70$ , considerando que valores em torno de .90 são considerados “excelentes”, ao redor de .80 são “muito bons” e em torno de .70 são “adequados” (Kline, 2015). Outro critério foi eliminar os itens de uma mesma dimensão que apresentaram correlações médias entre-itens  $< .40$  (Loiacono et al., 2002).

A fim de examinar as hipóteses realizámos ANCOVAs com três fatores 2 (agente de saúde) X 2 (decisão ética) X 2 (enquadramento da argumentação) para a aceitabilidade moral, responsabilidade moral, amabilidade, competência e confiabilidade, incluindo nos modelos as variáveis individuais e sociodemográficas que mostraram correlações significativas com as respetivas variáveis dependentes. Para examinar as variáveis dependentes relativas às questões exploratórias, nomeadamente a intenção de aceitar a medicação, a robô imaginada, e a robô adequada, realizámos análises de variância (ANOVAs), além de correlações lineares (*two-tailed*) com as variáveis dependentes principais. Utilizámos as orientações de Cohen (1992) para interpretar a dimensão do efeito das correlações lineares de Pearson para os valores de  $r$  (“baixo” ao redor de .10, “moderado” em torno de .30, e “elevado” ao redor de .50) e para a dimensão do efeito das ANCOVAs<sup>3</sup> ( $\eta_p^2$ ) com valores próximos a .01 sendo “baixos”, “moderados” em torno de .06, e “elevados” ao redor de .14.

---

<sup>3</sup> Algumas variáveis dependentes violaram o pressuposto da homogeneidade de variâncias. Porém, como a nossa amostra não apresenta uma dimensão pelo menos dupla em relação à menor amostra, o efeito

## 4. Resultados

### 4.1. Análises Preliminares

A fim de garantir a fiabilidade das variáveis dependentes, calculámos os respetivos alfas de Cronbach (ver Quadro 2). A escala usada para medir a aceitabilidade moral obteve consistência interna excelente (16 itens;  $\alpha = .90$ ) e a medida de responsabilidade moral atingiu uma boa consistência interna (3 itens;  $\alpha = .81$ ). Quanto às escalas utilizadas para medir os atributos percecionados na agente de saúde, a dimensão da amabilidade (3 itens) alcançou  $\alpha = .89$  após a exclusão do item 4 “bem-intencionada”, que apresentou correlação entre-itens  $< .40$ , e as dimensões da competência (6 itens;  $\alpha = .90$ ) e da confiabilidade (3 itens;  $\alpha = .78$ ) apresentaram consistência interna adequada.

As escalas usadas para medir as *variáveis individuais* apresentaram resultados algo heterogéneos. A escala das atitudes negativas em relação aos robôs (12 itens) alcançou boa consistência interna ( $\alpha = .88$ ). As dimensões dos traços morais relativas à relevância moral do cuidado (3 itens;  $\alpha = .79$ ) e da justiça (3 itens;  $\alpha = .78$ ) revelaram consistência interna adequada, no entanto foram excluídas as dimensões da autoridade (3 itens;  $\alpha = .62$ ) e as pertencentes à subescala do julgamento moral – cuidado (3 itens;  $\alpha = .38$ ), justiça (3 itens;  $\alpha = .19$ ), e autoridade (3 itens;  $\alpha = .54$ ) – pois não mostraram suficiente consistência interna. A dimensão do envolvimento com a saúde (16 itens;  $\alpha = .83$ ) mostrou boa consistência interna.

### 4.2. Análise Descritiva das Variáveis

Um resumo da estatística descritiva relativamente às variáveis sociodemográficas, individuais e dependentes analisadas neste estudo é apresentado no Quadro 2. Destacamos ainda que em relação à saúde 43.9% dos participantes afirmaram já ter prestado assistência de saúde domiciliar a alguém e 11.5% declararam já ter precisado deste tipo de assistência para si mesmos. A maioria relatou sua própria saúde geral como boa (60.1%) ou muito boa (19.8%).

---

da heterogeneidade das variâncias é pouco significativo (Guimarães & Cabral, 1997; Field, 2018), possibilitando a realização de análises de variância (ANOVAs) e ANCOVAs.

## Quadro 2

### Descrição Estatística das Variáveis

	N	Range	Min	Max	M	DP	$\alpha$ (Cronbach)
Idade	524	59	18	77	38.73	14.22	-
Escolaridade	501	5	1	6	3.53	1.03	-
Saúde Geral	499	4	1	5	2.02	0.67	-
PNARS	523	6	1	7	3.79	1.11	.88
Cuidado	518	5	1	6	5.04	0.80	.79
Justiça	517	5	1	6	5.15	0.76	.78
Envolvimento com a Saúde	502	3	2	5	3.83	0.48	.83
Aceitabilidade Moral	524	6	2	7	4.83	1.10	.90
Responsabilidade Moral	524	6	1	7	3.87	1.68	.81
Amabilidade	524	6	1	7	4.09	1.52	.89
Competência	524	6	1	7	4.93	1.24	.90
Confiabilidade	524	6	1	7	5.24	1.19	.78
IAM	524	6	1	7	5.14	1.52	-
Robô Imaginada	263	5	1	6	4	1.66	-
Robô Adequada	518	5	1	6	4.91	1.62	-

**Nota.** Atitudes negativas em relação aos robôs (PNARS); Intenção de Aceitar a Medicação (IAM)

Realizámos correlações entre as variáveis sociodemográficas e individuais e as variáveis dependentes (ver Quadro 3) com o intuito de identificar potenciais covariáveis a incluir em nossas análises principais. Estes resultados permitem ainda perceber o padrão da relação linear entre as variáveis.

## Quadro 3

### Correlações Lineares de Pearson ( $r$ ) entre as Variáveis Dependentes e as Potenciais Covariáveis

	Sex	Nac	Idade	Esc	Saúde	PNARS	Cuid	Just	Envolvimento com a Saúde
Aceitabilidade Moral	.01	-.03	.05	-.10*	.003	-.16**	.04	.05	.11*
Responsabilidade Moral	-.03	-.11*	-.08	-.12**	.07	.02	-.06	-.02	-.05
Amabilidade	-.07	-.07	-.03	-.06	.007	-.06	.10*	.06	.11*
Competência	.01	-.15**	-.11*	-.12**	-.04	-.17**	.02	.04	.14**
Confiabilidade	.04	-.03	-.06	-.10*	-.03	-.20**	.04	.10*	.12**

**Nota.** \* $p < .05$ , \*\* $p < .01$  (two-tailed). O sexo (Sex) foi codificado como 1 = Feminino e 2 = Masculino. A nacionalidade (Nac) foi codificada como 1 = Portuguesa e 2 = Brasileira. Os demais valores foram calculados utilizando as médias. Escolaridade (Esc); Atitudes negativas em relação aos robôs (PNARS); Cuidado (Cuid); Justiça (Just).

No que se refere às variáveis sociodemográficas, verificámos que os participantes portugueses ( $M = 4.12$ ,  $DP = 1.69$ ) atribuíram maior responsabilidade moral às agentes de saúde,  $t(493) = 2.50$ ,  $p = .013$ , do que os brasileiros ( $M = 3.73$ ,  $DP = 1.70$ ). Os portugueses ( $M = 5.19$ ,  $DP = 1.01$ ) também atribuíram mais competência às agentes,  $t(451) = 3.59$ ,  $p < .001$ , do que os brasileiros ( $M = 4.81$ ,  $DP = 1.33$ ). Uma maior escolaridade dos participantes se mostrou associada a um menor nível de aceitabilidade moral ( $r = -.10$ ,  $p = .032$ ) ao

comportamento da agente, a uma menor responsabilização moral ( $r = -.12, p = .008$ ) da agente, além de uma menor atribuição de competência ( $r = -.12, p = .006$ ) e de confiabilidade ( $r = -.10, p = .023$ ) à agente de saúde. Quanto maior a idade dos respondentes, menor a atribuição de competência ( $r = -.11, p = .014$ ) às agentes.

Quanto às variáveis individuais, verificamos que os participantes que reportaram atitudes menos negativas em relação aos robôs evidenciaram maior aceitabilidade moral ( $r = -.16, p < .001$ ) e maior percepção de competência ( $r = -.17, p < .001$ ) e de confiabilidade ( $r = -.20, p < .001$ ) em relação à agente e seu comportamento. Em termos de traços morais, maiores traços de cuidado ( $r = .10, p = .028$ ) estão associados à maior percepção de amabilidade, e maiores traços de justiça ( $r = .10, p = .031$ ) têm relação com uma maior atribuição de confiabilidade à agente. Os participantes que reportaram maior envolvimento com a própria saúde tendem a avaliar o comportamento da agente como sendo moralmente mais aceitável ( $r = .11, p = .014$ ) e a considerar a agente mais amável ( $r = .11, p = .019$ ), competente ( $r = .14, p = .002$ ) e confiável ( $r = .12, p = .010$ ).

### 4.3. Teste das Hipóteses em Função da Agente, da Decisão e da Argumentação

Considerando as correlações significativas encontradas entre as variáveis sociodemográficas e individuais e as variáveis dependentes, realizamos ANCOVAs com 3 fatores (ver Quadro 4) para cada uma das variáveis independentes<sup>4</sup> a fim de testar as nossas hipóteses e responder aos objetivos principais do estudo, controlando as possíveis covariáveis.

#### Quadro 4

*Análises de Covariância com 3 Fatores entre as Variáveis Independentes e as Variáveis Dependentes*

ANCOVAs	Agente de Saúde			Decisão Ética			Enquadramento da Argumentação		
	F	p	$\eta_p^2$	F	p	$\eta_p^2$	F	p	$\eta_p^2$
<b>Aceitabilidade Moral</b>	(1, 489) = 0.79	.375	0.002	(1, 489) = 43.91	<.001	0.082	(1, 489) = 0.002	.960	<0.001
<b>Responsabilidade Moral</b>	(1, 485) = 63.94	<.001	0.116	(1, 485) = 0.35	.554	0.001	(1, 485) = 0.86	.353	0.002
<b>Amabilidade</b>	(1, 492) = 6.78	.009	0.014	(1, 492) = 160.12	<.001	0.246	(1, 492) = 2.25	.135	0.005
<b>Competência</b>	(1, 481) = 0.34	.561	0.001	(1, 481) = 7.67	.006	0.016	(1, 481) = 0.05	.830	<0.001
<b>Confiabilidade</b>	(1, 487) = 0.00	.985	<0.001	(1, 487) = 6.46	.011	0.013	(1, 487) = 0.44	.509	0.001

**Notas.** Covariáveis incluídas: *Aceitabilidade Moral* - escolaridade, atitudes negativas em relação aos robôs (PNARS) e envolvimento com a saúde; *Responsabilidade Moral* - nacionalidade e escolaridade; *Amabilidade* - traço moral do cuidado e envolvimento com a saúde; *Competência* - nacionalidade, idade, escolaridade, PNARS e envolvimento com a saúde; *Confiabilidade* - escolaridade, PNARS, traço moral da justiça e envolvimento com a saúde.

<sup>4</sup> Observação: a agente de saúde foi codificada como -1 = humana e 1 = robô, a decisão ética como -1 = respeito à autonomia e 1 = beneficência/não-maleficência, e o enquadramento da argumentação como -1 = benefício à saúde e 1 = prejuízo à saúde.

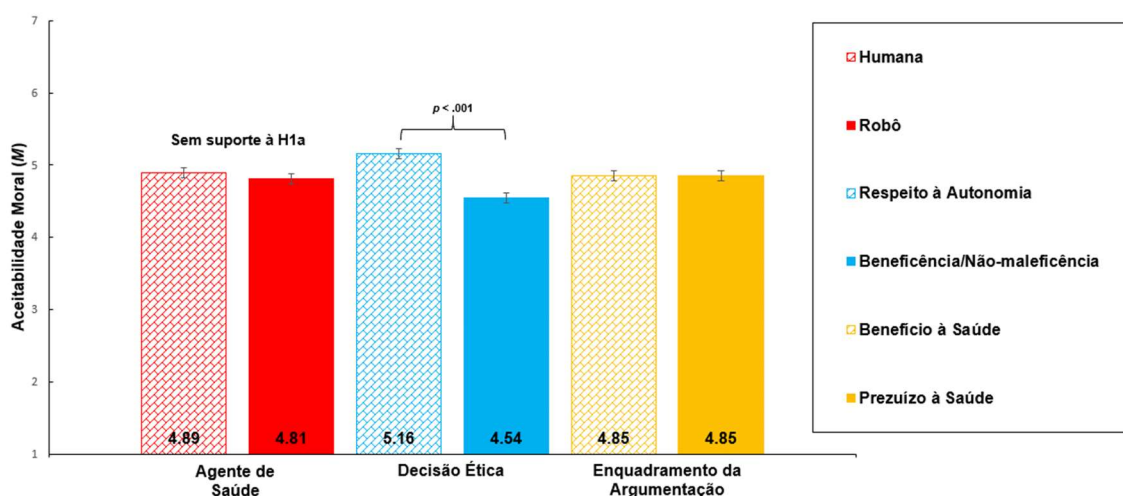
### 4.3.1. Aceitabilidade Moral do Comportamento da Agente de Saúde

Em relação à aceitabilidade moral atribuída ao comportamento da agente, a ANCOVA 2 (agente de saúde) x 2 (decisão ética) x 2 (enquadramento da argumentação) não mostrou efeitos significativos do tipo de agente,  $F(1, 489) = 0.788, p = .375, \eta_p^2 = 0.002$ , sem suporte à  $H_{1a}$ , nem da argumentação,  $F(1, 489) = 0.002, p = .960, \eta_p^2 = 0.000$ , mas verificamos um efeito principal significativo da decisão ética,  $F(1, 489) = 43.909, p < .001, \eta_p^2 = 0.082$  (Figura 2). Os participantes na condição em que a agente respeitou a autonomia da paciente relataram maior aceitabilidade moral da agente ( $M = 5.16, EP = 0.07$ ) do que os participantes na condição em que a agente escolheu o princípio da beneficência/não-maleficência ( $M = 4.54, EP = 0.07$ ).

As interações entre as variáveis independentes, inclusive entre a decisão ética e o tipo de agente, não apresentaram significância estatística ( $p_s > .05$ ), sugerindo que os efeitos da decisão ética foram semelhantes para os dois tipos de agente. De qualquer modo, para testar a  $H_2$ , analisamos os resultados das comparações simples entre o tipo de decisão ética em função de cada tipo de agente. Conforme prevíamos, os participantes que foram expostos à condição de agente robótica manifestaram maior aceitabilidade moral perante as robôs que respeitaram a autonomia ( $M = 5.07, EP = 0.09$ ) em comparação à condição em que as robôs optaram pela beneficência/não-maleficência ( $M = 4.54, EP = 0.09$ ),  $F(1, 489) = 16.332, p < .001, \eta_p^2 = 0.032$  (Figura 3). Estes resultados foram no mesmo sentido para as agentes humanas. Por fim, manteve-se estatisticamente significativo o efeito das covariáveis PNARS,  $F(1, 489) = 14.754, p < .001, \eta_p^2 = 0.029$ , escolaridade,  $F(1, 489) = 8.665, p = .003, \eta_p^2 = 0.017$ , e envolvimento com a saúde,  $F(1, 489) = 8.069, p = .006, \eta_p^2 = 0.015$ , na aceitabilidade moral da agente.

Figura 2

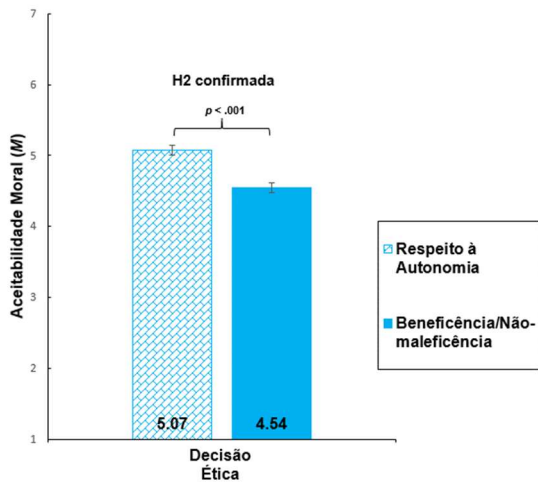
Aceitabilidade Moral do Comportamento da Agente de Saúde





**Figura 3**

*Aceitabilidade Moral das Agentes Robôs pelo Tipo de Decisão Ética*

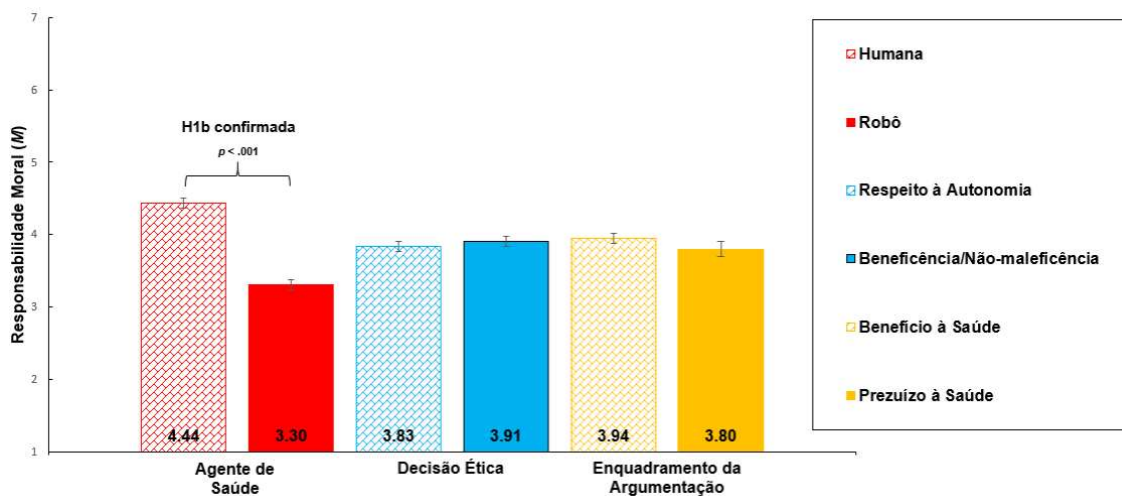


### 4.3.2. Responsabilidade Moral da Agente de Saúde

A ANCOVA 2x2x2 não mostrou efeitos significativos da decisão ética,  $F(1, 485) = 0.35, p = .554, \eta_p^2 = 0.001$ , e da argumentação,  $F(1, 485) = 0.86, p = .353, \eta_p^2 = 0.001$ , mas ocorreu efeito principal significativo do tipo de agente na responsabilidade moral,  $F(1, 485) = 63.94, p < .001, \eta_p^2 = 0.12$ . A responsabilidade moral atribuída pelos participantes à agente humana ( $M = 4.44, EP = 0.10$ ) foi maior do que à agente robótica ( $M = 3.30, EP = 0.10$ ), com suporte à  $H_{1b}$  (Figura 4). Não verificamos interações ( $p_s > .05$ ) entre as variáveis independentes, mas foram mantidos os efeitos significativos das covariáveis escolaridade,  $F(1, 485) = 5.90, p = .015, \eta_p^2 = 0.01$ , e nacionalidade,  $F(1, 485) = 4.83, p = .028, \eta_p^2 = 0.01$ , na responsabilidade moral.

**Figura 4**

*Responsabilidade Moral da Agente de Saúde*

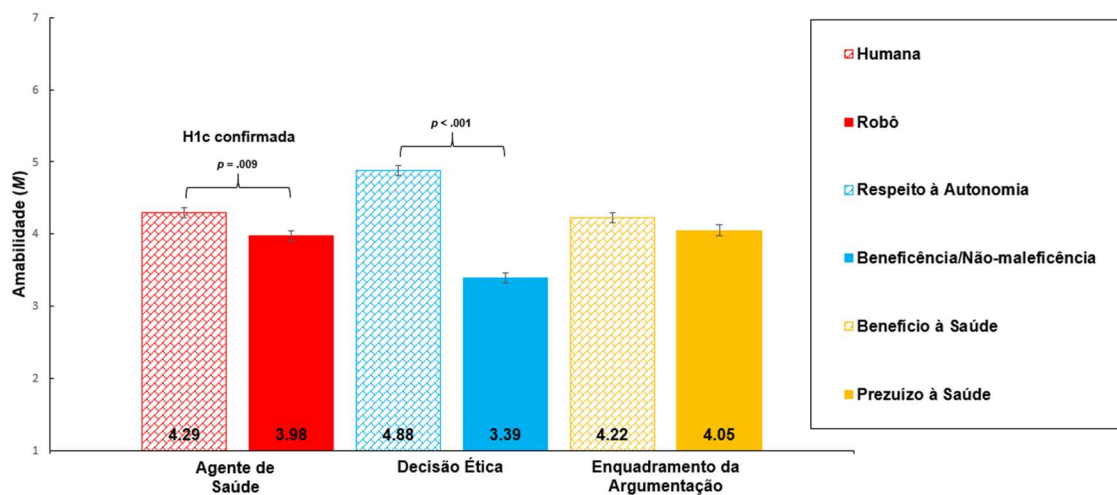


### 4.3.3. Avaliação dos Atributos da Agente de Saúde

Em relação à avaliação da amabilidade da agente de saúde, a ANCOVA com os mesmos três fatores mostrou efeitos significativos do tipo de agente,  $F(1, 492) = 6.78, p = .009, \eta_p^2 = 0.01$ , e da decisão ética,  $F(1, 492) = 160.12, p < .001, \eta_p^2 = 0.25$ , mas nenhum efeito da argumentação,  $F(1, 492) = 2.25, p = .135, \eta_p^2 = 0.005$ . Os participantes perceberam a agente humana como sendo um pouco mais amável ( $M = 4.28, EP = 0.08$ ) do que a agente robótica ( $M = 3.98, EP = 0.08$ ), com suporte à  $H_{1c}$ . Além disso, a agente que decidiu respeitar a autonomia foi avaliada como sendo muito mais amável ( $M = 4.88, EP = 0.08$ ) do que a agente que optou pela beneficência/não-maleficência ( $M = 3.39, EP = 0.08$ ) (Figura 5). Não houve interações significativas entre as variáveis independentes ( $p_s > .05$ ). O efeito do envolvimento com a saúde manteve-se estatisticamente significativo,  $F(1, 492) = 6.69, p = .010, \eta_p^2 = 0.01$ , na avaliação da amabilidade da agente.

**Figura 5**

*Percepção de Amabilidade na Agente de Saúde*

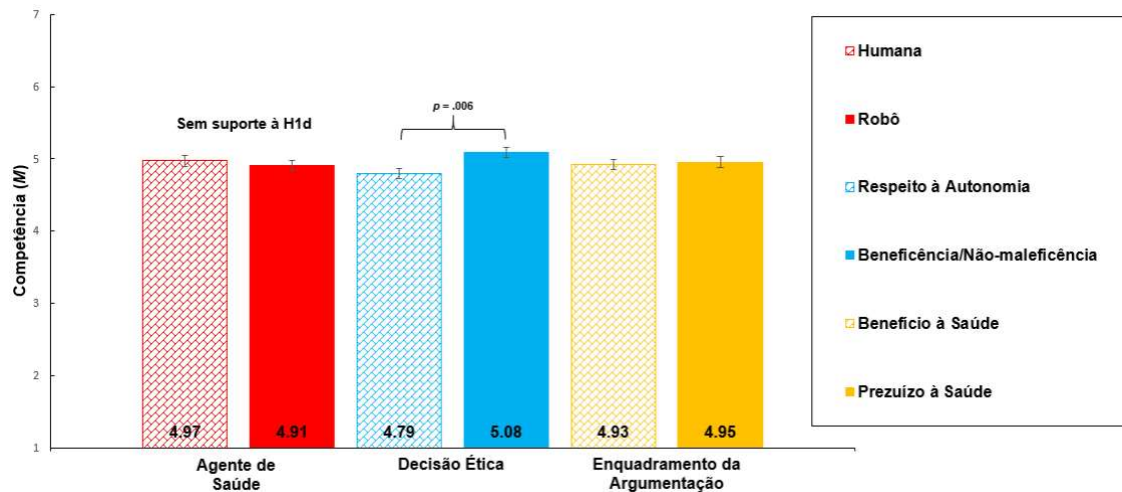


Em relação à avaliação da competência da agente de saúde, a ANCOVA 2x2x2 não revelou efeitos significativos da argumentação,  $F(1, 481) = 0.05, p = .830, \eta_p^2 = 0.00$ , ou do tipo de agente,  $F(1, 481) = 0.34, p = .561; \eta_p^2 = 0.001$ , sem suporte à  $H_{1d}$ , porém ocorreu um efeito principal significativo do tipo de decisão,  $F(1, 481) = 7.67, p = .006, \eta_p^2 = 0.02$ . A opção pela beneficência/não-maleficência ( $M = 5.08, EP = 0.07$ ) levou a uma maior atribuição de competência à agente do que o respeito à autonomia ( $M = 4.79, EP = 0.07$ ) (Figura 6). Não houve interações significativas ( $p_s > .05$ ) entre as variáveis independentes, mas os efeitos significativos da PNARS,  $F(1, 481) = 24.25, p < .001, \eta_p^2 = 0.05$ , do envolvimento com a saúde,

$F(1, 481) = 13.41, p < .001, \eta_p^2 = 0.03$ , e da nacionalidade,  $F(1, 481) = 7.27, p = .007, \eta_p^2 = 0.015$ , na competência foram mantidos.

**Figura 6**

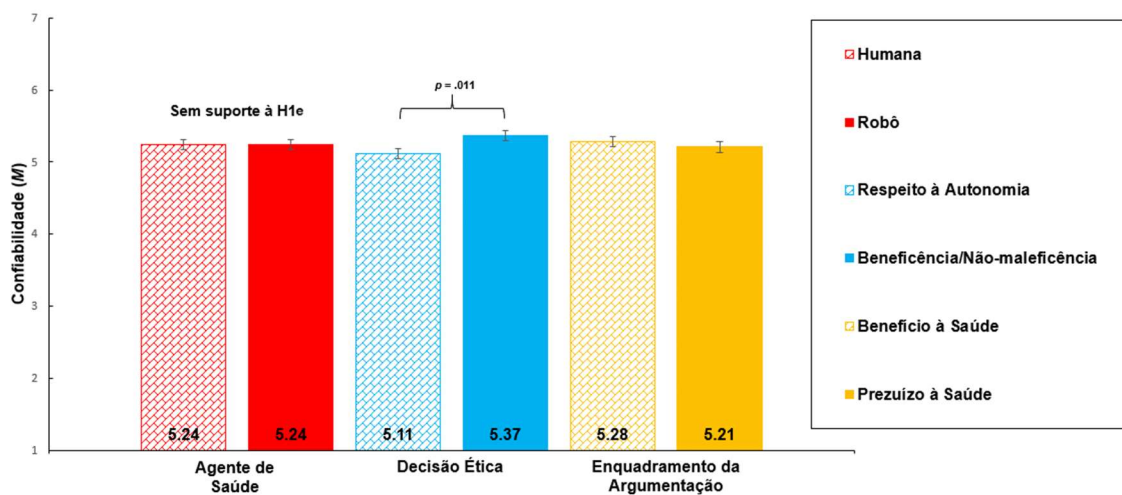
*Percepção de Competência na Agente de Saúde*



A ANCOVA 2x2x2 mostrou um efeito principal significativo na confiabilidade como consequência da decisão ética,  $F(1, 487) = 6.46, p = .011, \eta_p^2 = 0.01$ . A opção pela beneficência/não-maleficência ( $M = 5.37, EP = 0.07$ ) levou a uma maior percepção de confiabilidade em relação à agente do que o respeito à autonomia ( $M = 5.11, EP = 0.07$ ). A argumentação,  $F(1, 487) = 0.438, p = .509, \eta_p^2 = 0.001$ , e o tipo de agente,  $F(1, 487) = 0.00, p = .985, \eta_p^2 = 0.00$ , não mostraram efeitos significativos, *sem suporte à H<sub>1e</sub>* (Figura 7).

**Figura 7**

*Percepção de Confiabilidade na Agente de Saúde*

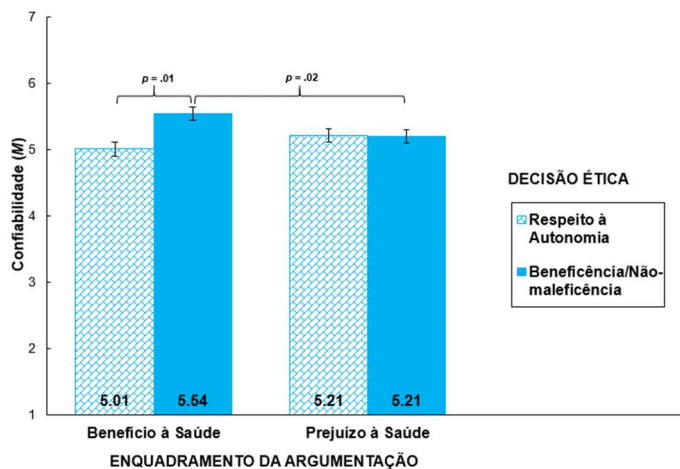


Verificamos uma interação entre a decisão ética e o enquadramento da argumentação,  $F(1, 487) = 6.77, p = .010, \eta_p^2 = 0.01$ , com efeito significativo na confiabilidade. Mediante testes adicionais verificamos que a confiabilidade foi superior quando a agente decidiu pela beneficência/não-maleficência junto à argumentação do benefício à saúde ( $M = 5.54, EP = 0.10$ ), comparativamente à decisão ética pelo respeito à autonomia aliada ao argumento do benefício à saúde ( $M = 5.01, EP = 0.11$ ),  $F(1, 487) = 12.84, p < .001, \eta_p^2 = 0.03$ , e à decisão ética pela beneficência/não-maleficência com o argumento do prejuízo à saúde ( $M = 5.21, EP = 0.10$ ),  $F(1, 487) = 5.49, p = .020, \eta_p^2 = 0.01$  (Figura 8). As demais interações entre as variáveis independentes não apresentaram efeitos significativos ( $p_s > .05$ ).

Por último, foram mantidos os efeitos significativos da PNARS,  $F(1, 487) = 22.46, p < .001, \eta_p^2 = 0.04$ , escolaridade,  $F(1, 487) = 9.95, p = .002, \eta_p^2 = 0.02$ , envolvimento com a saúde  $F(1, 487) = 7.06, p = .008, \eta_p^2 = 0.01$ , e justiça,  $F(1, 487) = 6.37, p = .012, \eta_p^2 = 0.01$ , na confiabilidade.

**Figura 8**

*Confiabilidade na Interação entre a Decisão Ética e a Argumentação*



## 4.2. Análises Exploratórias Adicionais

### 4.2.1. Intenção de Aceitar a Medicação

A ANOVA 2 (agente de saúde) x 2 (decisão ética) x 2 (enquadramento da argumentação) mostrou que a autoavaliação do nível de intenção dos participantes de aceitar a recomendação da agente e tomar a medicação não apresentou diferenças para nenhuma das variáveis independentes ( $p_s > .05$ ). O efeito de persuasão do argumento do benefício à saúde foi semelhante ao do prejuízo à saúde em todas as condições expostas aos participantes.

Entretanto, a intenção de aceitar a medicação apresentou correlações significativas ( $p_s < .001$ ) com algumas variáveis dependentes, designadamente a aceitabilidade moral ( $r = .35$ ), a competência ( $r = .31$ ), a amabilidade ( $r = .30$ ) e a confiabilidade ( $r = .25$ ).

#### 4.2.2. Robô Imaginada e Robô Adequada

A escala de semelhança humana com as fotos de robôs não mostrou efeitos na ANOVA 2 (decisão ética) x 2 (enquadramento da argumentação) para a *robô imaginada* ( $p_s > .05$ ), mas verificámos correlações significativas da robô imaginada com algumas variáveis dependentes. Quanto maior o índice de semelhança humana da robô imaginada, mais competência ( $r = .24$ ,  $p < .001$ ), amabilidade ( $r = .14$ ,  $p = .02$ ) e confiabilidade ( $r = .14$ ,  $p = .03$ ) foram atribuídas à agente robótica, além de maior aceite moral ( $r = .12$ ,  $p = .05$ ) ao comportamento da agente.

Mais do que 50% dos participantes na condição de agente robótica escolheram as fotos de robôs com índice de semelhança humana mais elevados, icub (26.2%) e nadine (24.0%), enquanto mais próximas das robôs que tinham imaginado ao ler a vinheta do dilema moral.

A ANOVA 2 (agente de saúde) x 2 (decisão ética) x 2 (enquadramento da argumentação) para a *robô adequada* não revelou efeitos para a decisão e para a argumentação ( $p_s > .05$ ), mas verificámos efeito significativo para o tipo de agente,  $F(1, 510) = 6.55$ ,  $p = .01$ ,  $\eta_p^2 = 0.01$ . A semelhança humana relativa à robô considerada adequada pelos participantes na condição de agente robô ( $M = 5.10$ ,  $EP = 0.10$ ) foi significativamente maior do que na condição de agente humana ( $M = 4.73$ ,  $EP = 0.10$ ). Não ocorreram correlações significativas entre a avaliação da adequação da robô e as variáveis dependentes ( $p_s > .05$ ).

A maioria dos participantes escolheu como robôs mais adequadas à enfermagem as de maior semelhança humana, nadine (57.3%) e icub (18.9%), que em conjunto correspondem a 75.4% das escolhas.

#### 4.2.3. Compreensão e Credibilidade

Somente 3.5% dos participantes alegaram discordar ao menos um pouco (nenhum participante discordou totalmente) relativamente à boa compreensão da situação descrita na vinheta, enquanto 81.9% indicaram ter compreendido bem. Alegaram discordar ao menos um pouco do realismo da situação descrita pela vinheta 10.2% dos participantes, enquanto 67.1% concordaram que a vinheta pareceu bastante realista, e 17.0% concordaram um pouco. Por outras palavras, 84.1% dos participantes concordaram que a vinheta foi realista.



## 5. Discussão

Mediante uma abordagem experimental, procurámos contribuir para a compreensão dos julgamentos morais relativos aos dilemas morais no contexto da HRI em ambientes de cuidados de saúde. Recolhemos dados de portugueses e brasileiros através de questionários online e analisámos os julgamentos morais dos participantes acerca do respeito à autonomia vs. a ênfase à beneficência/não maleficência em relação a uma paciente que rejeita um medicamento, numa decisão tomada por agentes (robô vs. humana) que utilizaram de diferentes argumentos (benefício à saúde vs. prejuízo à saúde) referentes à importância de aderir à medicação. Além disso, os participantes avaliaram o carácter moral da agente de saúde em termos dos atributos da amabilidade, da competência, e da confiabilidade.

### 5.1. Aceitabilidade Moral enquanto Julgamento de Erro Moral

A dimensão da aceitabilidade moral enquadra-se nos julgamentos de erro moral, que estabelecem se uma determinada ação (ou inação), em geral já concretizada, foi intencional e violou uma ou mais normas éticas (Malle, 2021). Considera ainda a existência ou não de justificativas pertinentes, além das possíveis consequências do ato (Voiklis et al., 2016).

Os resultados da nossa medida de julgamento de erro moral foram definidos pelo tipo de decisão ética, sendo que a aceitabilidade moral não apresentou diferenças significativas entre as agentes, sem confirmação para a  $H_{1a}$ . Por outro lado, independentemente do tipo de agente (humana ou robô), verificou-se no presente estudo que os participantes expostos à decisão ética de respeito à autonomia julgaram esta decisão como sendo significativamente mais aceitável moralmente do que os participantes na condição da priorização à beneficência/não-maleficência (opção “utilitária”), confirmando a  $H_2$ .

Estes resultados são divergentes aos obtidos por Malle et al. (2015), estudo no qual um robô que sacrifica uma pessoa para salvar quatro (opção “utilitária”) teve a sua ação avaliada como moralmente mais aceitável do que a sua inação, ao contrário do agente humano, para o qual a aceitabilidade moral foi muito superior na condição de inação quando comparada com a do sacrifício.

Em Malle et al. (2016), os resultados na avaliação da aceitabilidade moral foram similares aos nossos quanto ao tipo de agente, sem diferenças significativas entre o robô mecânico, o robô humanoide, a IA, e o humano, mas mostrou resultados divergentes quanto à decisão ética, sem diferenças significativas na aceitabilidade moral entre o sacrifício e a inação. Salientamos que nossas medidas têm metodologia diferente da utilizada em ambos os estudos de Malle et

al. (2015, 2016), na medida em que avaliaram o erro moral por meio de um item com resposta binária.

Por contraste, os nossos resultados relativos à decisão ética são consistentes com os de Laakasuo et al. (2019), investigação na qual a maior aceitabilidade moral atribuída ao respeito à autonomia (vs. a beneficência/não-maleficência) não variou significativamente consoante o tipo de agente, embora os robôs tenham sido julgados como sendo significativamente menos aceitáveis moralmente do que os humanos ao medicar o paciente à força. Contudo, o nosso dilema moral tem o atenuante de não incluir a medicação forçada, mas apenas a negação da vontade da paciente enquanto opção “utilitária”. Talvez esta diferença tenha contribuído para que ocorresse uma aversão dos participantes aos robôs que manifestaram comportamento coercivo (inexistente em nosso estudo) perante um humano nos dilemas dos estudos de Laakasuo et al. (2019), implicando num maior julgamento de erro moral. Ao respeitar a autonomia do paciente nos estudos de Laakasuo e colaboradores, o agente de saúde desobedece às instruções do médico responsável (humano ou IA), configurando uma descrição de desobediência inexistente em nossas vinhetas, embora esta distinção não pareça ter originado diferenças perceptíveis entre as investigações. Ressaltamos que a nossa medida da aceitabilidade moral é similar à utilizada por Laakasuo e colaboradores por ter sido adaptada a partir do seu trabalho.

Na avaliação dos efeitos da manipulação na *aceitabilidade moral* foram ainda consideradas algumas covariáveis. Os resultados mostraram que esta avaliação da agente (robô ou humana) e de seu comportamento mostrou estar associada às atitudes negativas em relação aos robôs, sugerindo que quanto menos negativas forem as atitudes dos participantes face aos robôs, maior a avaliação da aceitabilidade moral. Além disso, um menor grau de escolaridade dos participantes e um maior nível de envolvimento com a própria saúde também se mostraram associadas a uma maior aceitabilidade moral das agentes e do seu comportamento. É digno de nota que todas as vinhetas faziam referência a um sistema de IA na instrução médica a ambas as agentes, o que pode ter contribuído para que ocorresse esta relação entre as atitudes negativas em relação aos robôs e a avaliação da aceitabilidade moral, dada a proximidade conceitual entre os robôs e a IA. Além disso, o maior envolvimento das pessoas com a própria saúde talvez dê origem a uma atitude mais positiva perante às novas tecnologias de saúde. Em contrapartida, um maior nível de escolaridade parece contribuir para uma postura mais crítica das pessoas relativamente à aceitabilidade moral.



## 5.2. Responsabilidade Moral enquanto Julgamento de Culpa

A dimensão da responsabilidade moral tem foco no nível de responsabilização imputado pelos participantes à agente de saúde acerca de sua decisão e suas possíveis consequências, além da culpabilização decorrente destas consequências. A responsabilidade moral é análoga aos julgamentos de culpa realizados pelas pessoas, que se fundamentam nos julgamentos de erro moral, mas também numa avaliação mais profunda das justificativas e das intenções do agente (Malle, 2021), além da capacidade de escolha e de agência mental atribuídas ao agente (Voiklis et al., 2016).

De acordo com os nossos resultados, a responsabilidade moral atribuída pelos participantes à agente humana foi significativamente maior do que à agente robô em todas condições, confirmando a  $H_{1b}$ .

Este nosso resultado diverge novamente do estudo de Malle et al. (2015), que mostrou: (a) a culpabilização superior do agente humano na opção pelo sacrifício (vs. a inação), (b) a culpa atribuída ao robô sem diferenças significativas relativamente à decisão ética, e (c) uma atribuição global de culpa moral quantitativamente semelhante a ambos os agentes.

As três experiências do estudo de Malle et al. (2016) também apresentaram resultados algo diferentes dos nossos, mas em geral os agentes (humano, robô mecânico, robô humanoide, e a IA) foram significativamente mais culpabilizados pela decisão do sacrifício. Somente na primeira experiência os robôs mecânicos foram mais culpados pela inação (com resultado oposto para os demais agentes), quando o texto do dilema não enfatizou a vítima (a pessoa a ser sacrificada).

De qualquer modo, Malle et al. (2015, 2016) utilizaram questões e medidas diferentes das nossas, com um item avaliado através de uma barra deslizante, mediante o qual a dimensão de culpa parece ter sido avaliada de modo mais automático e condenatório, relacionada ao sacrifício de um ser humano. A nossa dimensão da responsabilidade moral parece ter propiciado um julgamento de culpa mais aprofundado e sistemático acerca de um dilema moral menos contundente, porém mais complexo e detalhado do que o dilema do trolley adaptado aos robôs. Por exemplo, Malle et al. (2019) mostraram diferenças nos julgamentos de culpa entre agentes artificiais (IAs e *drones*) e humanos que podem ter sido originadas de percepções distintas acerca dos papéis associados a ambos e às justificativas morais relativas a estes papéis. Os participantes do estudo de Malle et al. (2019) não forneceram explicações com base em papéis para seus julgamentos morais aos agentes artificiais, ocorrendo o contrário aos agentes humanos, para os quais as avaliações morais foram moduladas por seus papéis sociais ou institucionais. Da mesma forma, a narrativa com a agente de saúde no nosso dilema pode ter

oferecido um contexto social mais detalhado que possibilitou aos participantes um julgamento de culpa mais informado.

O nosso resultado acerca da responsabilização moral reproduziu os resultados de Laakasuo et al. (2019), que apresentaram índices de responsabilidade moral significativamente superiores para o humano em relação ao robô, independentemente de suas decisões éticas. Ressaltamos que Laakasuo e colaboradores utilizaram uma medida com um item e nós utilizámos uma compósita de três itens. Os nossos resultados também são consistentes com os de Kahn et al. (2012), estudo no qual a maioria dos participantes responsabilizou moralmente um agente humano em maior grau que um robô social que causou prejuízo num jogo.

De facto, a responsabilização moral de um robô é discutível, considerando que é um agente artificial sem uma *mente* (ToM - *Theory of Mind*: Premack & Woodruff, 1978; Carruthers & Smith, 1998; Leslie et al., 2004) ou uma consciência humana (de Graaf & Ben Allouch, 2016; Devillers, 2021). Dodig-Crnkovic e Çürüklü (2012) propõem a atribuição de responsabilidades funcionais aos agentes morais artificiais numa rede de responsabilidade dividida com outros intervenientes num sistema sócio-tecnológico. Esta proposta não nos parece adequada, sobretudo devido à diferença ontológica evidente entre humanos e agentes artificiais. Por exemplo, Sharkey (2017) defende que os seres humanos sempre serão moralmente responsáveis pelas consequências das ações engendradas por robôs. O facto de um agente deter um papel e ser encarregue de uma determinada tarefa o remete a expectativas morais e normativas de outros indivíduos, e a fiabilidade de ser considerado moralmente responsável é parte inextricável de seus relacionamentos sociais (Williams, 2012; Järvik, 2013). A capacidade de um agente cumprir normas morais é avaliada em comparação aos outros, e determinamos se uma pessoa nos é semelhante para então imaginar como nós poderíamos ter atuado numa determinada situação (Järvik, 2013). Aristóteles argumentou que um agente só pode ser responsabilizado moralmente se tem consciência do que está a fazer, se sua ação é voluntária e sua decisão, autónoma (Järvik, 2013; Eshleman, 2014; Beauchamp & Childress, 2019), o que nos parece estar bem longe das possibilidades de um agente artificial.

Em relação às covariáveis, verificou-se que os participantes com maior grau de escolaridade imputaram menor responsabilidade moral à agente de saúde. A escolaridade de maior nível parece contribuir para uma perspectiva crítica que enfraquece a responsabilização à agente, mas o motivo para isto merece novas investigações. Por outro lado, os portugueses responsabilizaram moralmente mais as agentes do que os brasileiros. Estes resultados vão ao encontro das evidências de que os brasileiros são mais tolerantes relativamente à obediência a regras e à evitação da incerteza do que os portugueses (Hofstede, 2011).

### 5.3. Avaliação dos Atributos da Agente enquanto Julgamentos de Caráter Moral

A avaliação dos traços da agente em termos da amabilidade, competência e confiabilidade podem ser enquadradas como julgamentos de caráter moral baseados em eventos e consistem numa forma de avaliação da personalidade (Malle, 2021).

Conforme esperávamos, a  $H_{1c}$  foi confirmada na medida em que as agentes humanas foram avaliadas pelos participantes como significativamente mais amáveis que as robôs. Por outro lado, os traços de competência e de confiabilidade das agentes não foram percebidas de modo distinto, sem confirmação de  $H_{1d}$  e  $H_{1e}$ . Os participantes atribuíram às agentes amabilidade muito maior e menos competência e confiabilidade nas condições do respeito à autonomia da paciente por comparação com a opção pela beneficência/não-maleficência (opção “utilitária”). A condição da decisão pela beneficência/não-maleficência, ao contrário, gerou a percepção de agentes um pouco mais confiáveis e competentes, mas muito menos amáveis, do que a opção pelo respeito à autonomia.

Estes resultados são consistentes com de Graaf e Ben Allouch (2016), cujos participantes relataram a falta de competências sociais e emocionais aos robôs e, especificamente na área dos cuidados de saúde, apontaram um possível aumento da eficiência nas tarefas executadas por agentes artificiais às custas da impessoalidade no trato com os pacientes. A competência e a confiabilidade, em seus significados corriqueiros, podem ser atribuídas a qualquer tipo de máquina e estão ligadas a aspetos racionais e de desempenho utilitário, enquanto que a amabilidade está associada a aspetos afetivos de atribuição incomum a um agente artificial. As pessoas parecem utilizar os mesmos conceitos e explicações para o comportamento de agentes humanos e robôs, com robustas inferências de intencionalidade e de mente, mas de um modo um pouco diferente entre ambos, sendo que os robôs tendem a ser percebidos em geral como seres racionais em vez de afetivamente motivados (de Graaf & Malle, 2019).

Os nossos resultados são parcialmente análogos aos de Malle et al. (2016), no qual uma compósita das dimensões da confiança, inteligência, e *well-liked* (popularidade), nomeada avaliação social, para os agentes artificiais e humanos no dilema do trolley. No estudo de Malle et al. (2016), os agentes que decidiram pelo sacrifício (opção “utilitária”) foram avaliados mais positivamente<sup>5</sup>, mas os humanos foram avaliados mais positivamente do que qualquer agente artificial. No nosso estudo, os atributos da agente variaram em função da sua decisão ética, e a amabilidade foi influenciada também pelo tipo de agente, mas com uma dimensão de efeito

---

<sup>5</sup> É possível que a “avaliação social” de Malle et al. (2016) tenha correlação positiva com a opção “utilitária” devido à prevalência de itens associados às dimensões da competência (1 item) e da confiabilidade (2 itens) contra apenas 1 item da amabilidade.

menor em relação à dimensão de efeito da decisão ética. Da mesma forma, a avaliação social de Malle e colaboradores sofreu influência tanto da decisão ética como do tipo de agente, mas com uma dimensão de efeito menor para o tipo de agente em relação à dimensão de efeito da decisão ética. Podemos concluir que a decisão ética parece ter um peso maior do que o tipo de agente para os participantes nos seus julgamentos de caráter moral.

Gleichgerrcht e Young (2013) encontraram no dilema do trolley uma relação significativa e específica entre os julgamentos morais “utilitários” (opção pelo sacrifício) dos participantes com seus próprios baixos índices de preocupação empática. O calor humano face a alguém em dificuldades é uma característica de quem demonstra preocupação empática, que pode surgir quando alguém é vitimizado. No nosso dilema moral foi apresentada uma condição menos extrema (não há fatalidade envolvida na escolha), sendo que a nossa vinheta mostra uma decisão ética da agente e a avaliação é sobre a amabilidade relativa a esta agente. O julgamento “utilitário” no estudo de Gleichgerrcht e Young foi realizado pelos participantes e a avaliação da preocupação empática foi relativa a eles mesmos. Entretanto, as dimensões relativas ao calor humano (respetivamente, a amabilidade e a preocupação empática) apresentaram correlação negativa com as decisões “utilitárias” tanto no nosso estudo como no de Gleichgerrcht e Young, ainda que em âmbitos diferentes.

Laakasuo et al. (2019) verificaram que os agentes humanos foram avaliados pelos participantes como mais confiáveis do que os robôs e não ocorreu variação da confiabilidade em função da decisão ética. Este resultado diverge do nosso relativamente à confiabilidade, que variou em decorrência da decisão ética e não consoante o tipo de agente. No entanto, os itens de Laakasuo e colaboradores versavam sobre confiar a própria vida ou a de familiares ao agente de saúde, o que talvez tenha agregado muito afeto à dimensão da confiabilidade, enviesando o julgamento de caráter moral no sentido do calor humano. Isto pode explicar a variação da confiabilidade consoante o tipo de agente de uma forma equivalente às evidências de nosso estudo quanto à amabilidade, e pode ter evitado possíveis variações da confiabilidade em função da decisão ética no estudo de Laakasuo et al. (2019).

Gamez et al. (2020) sugerem que a teoria ética da virtude parece promissora para o desenvolvimento e treinamento de agentes artificiais, embora as justificativas das pessoas para a atribuição de virtudes (caráter moral) a estes agentes sejam mais complexas e variadas, além de mais fracas, em comparação com os agentes humanos. As inferências utilizadas nos julgamentos de caráter moral são sensíveis às justificativas e motivações dos comportamentos do agente (Martin & Cushman, 2016; Reeder et al., 2002), e as pessoas geralmente integram os dados sobre o caráter através do aumento das inferências relativas ao estado mental do agente

(Kim et al. 2020). Considerando que uma agente robótica não tem mente (ToM) nem consciência humana e não pode – ao menos no momento atual (Devillers, 2021) – apresentar autoconsciência ou autocrítica, a ligação entre o “caráter moral” de um robô e o julgamento de suas ações éticas atribuídas por um observador parece discutível. Se uma agente robótica não possui uma mente ou consciência humana e possui um estatuto ontológico diferenciado (Kahn, 2013; Kim, 2019; Sommer, 2019), a quem é possível atribuir as suas características morais? Os estudos de Kahn et al. (2011, 2012) dão suporte à hipótese de uma Nova Categoria Ontológica, ou seja, um novo modo de perceber e distinguir um grupo de agentes – neste caso, os robôs corporificados. Os participantes da investigação de Kahn et al. (2012) interagiram com um robô social, sendo que 47.5% afirmou tratar-se de uma tecnologia e 52.5% o descreveu como algo entre uma tecnologia e um ser vivo. A maioria acreditava que o robô podia pensar, 50% afirmou que ele estava consciente, 63% atribuiu-lhe confiabilidade, mas somente 35% relatou crer que o robô tinha sentimentos. Muitos argumentaram que o robô poderia pensar ou ter emoções, mas de modo qualitativamente diferente de um ser humano. No nosso estudo, os participantes julgaram as agentes humanas e robôs de modo similar quanto ao caráter moral, exceto em termos de amabilidade, o que parece evidenciar uma tendência das pessoas a imputar uma “personalidade” aos robôs, mas de um modo um pouco diferente do que a um ser humano.

Em relação às covariáveis, verificou-se que os participantes com maior nível de envolvimento com a saúde perceberam mais amabilidade nas agentes, talvez devido à explicação oferecida pelas agentes acerca da necessidade da medicação. As pessoas mais envolvidas com a própria saúde possivelmente apreciam estas informações, sendo que investigações futuras podem ser importantes no sentido de verificar esta hipótese.

Quanto menos atitudes negativas perante os robôs foram reportadas pelos participantes, maior foi a atribuição de competência às agentes, possivelmente por causa de uma maior receptividade às novas tecnologias na saúde (e.g., os robôs e a IA). Os participantes que relataram mais envolvimento com a própria saúde também imputaram maior competência às agentes, talvez pelo facto de terem oferecido uma explicação acerca do medicamento. Por outro lado, os portugueses atribuíram mais competência às agentes do que os brasileiros, resultado inesperado e sem explicação, que merece investigação relativamente a possíveis diferenças culturais.

Verificou-se ainda que uma maior confiabilidade foi atribuída às agentes pelos participantes que revelaram atitudes menos negativas frente aos robôs, talvez por sua maior transigência às novas tecnologias. O reporte de maior envolvimento do participante com a sua própria saúde acompanhou uma maior atribuição de confiabilidade à agente, o que pode estar relacionado com um maior conhecimento das novas tecnologias aplicadas à saúde. Um menor

grau de escolaridade dos participantes e traços morais de justiça mais elevados também acompanharam uma maior confiabilidade atribuída às agentes. Embora o presente projeto não tivesse o objetivo de analisar a relação entre estas variáveis e as nossas variáveis dependentes, o facto de terem apresentado resultados significativos mostram a sua importância em considerá-las em estudos subsequentes e aprofundar a sua análise.

### **5.3.1. Interação entre a Decisão e a Argumentação na Confiabilidade**

Praticamente toda a variação da confiabilidade ocorreu na interação da decisão pela beneficência/não-maleficência com o argumento do benefício à saúde. Mas por que motivo os participantes consideraram mais confiável a agente de saúde nestas condições?

Cameron et al. (2015) defendem duas dimensões a ser examinadas ao analisar as emoções nos julgamentos morais: o afeto central (*valência e arousal*) e o conhecimento conceitual acerca do contexto (tipo e local da ação, tipos de agente e paciente morais, cultura, e outros aspetos específicos). Considerando o conteúdo conceitual associado à interação entre a beneficência/não-maleficência e o benefício à saúde, observamos uma coerência entre a decisão da agente e a sua argumentação. A agente mostrou-se envolvida no seu ofício e apresentou argumentos, o que pode ter afetado positivamente a avaliação dos participantes quanto à sua confiabilidade, dado que a percepção do caráter moral é sensível às motivações e justificativas da agente (Martin & Cushman, 2016; Reeder et al., 2002). Além disso, o argumento do benefício à saúde pode ter eliciado a esperança (Nabi et al., 2020), emoção de valência positiva que pode funcionar como uma sinalização de incentivo (Damásio, 2019) e reforçar o conteúdo conceitual (Cameron et al., 2015) associado à confiabilidade.

As interações onde a decisão ética foi respeitar a autonomia da paciente não mostram a mesma consistência de argumentação, pois esta é dada no que respeita à importância da medicação – o porquê de a autonomia ter sido respeitada não é explicado pela agente de saúde. Resta então a vineta na qual a agente privilegia a beneficência/não-maleficência à paciente sob o argumento de prejuízo à saúde. Neste caso, apesar de a agente ter apresentado justificativas para sua decisão, o efeito da interação na confiabilidade foi mais negativo do que nas outras condições. Talvez o argumento do prejuízo à saúde tenha causado medo (Nabi et al., 2020), emoção de valência negativa que geralmente expressa um alarme (Damásio, 2019) passível de suscitar uma conduta de evitação (Cameron et al., 2015), o que pode ter afetado negativamente a confiabilidade nesta condição.

#### **5.4. Intenção de Aceitar a Medicação**

No presente estudo, não houve diferenças significativas entre os dois tipos de enquadramento da argumentação (benefício à saúde vs. prejuízo à saúde) no que tange à intenção dos participantes em aceitar a recomendação da agente de saúde. Este resultado é coerente com as meta-análises anteriormente citadas (Gallagher & Updegraff, 2011; Nabi et al., 2020; O’Keefe & Jensen, 2007). Entretanto, os participantes que relataram maior intenção de aceitar a medicação atribuíram à agente mais amabilidade e competência, o que talvez tenha afetado positivamente a intenção de aceitar a medicação através de pistas heurísticas (e.g., a perícia na saúde, a atratividade e a simpatia; Lerner et al., 2015) percebidas na agente de saúde. Os nossos resultados são consistentes com o estudo de Dai e MacDorman (2018), no qual uma maior percepção de amabilidade e de competência em relação ao agente de saúde contribuiu para aumentar a intenção de adesão dos participantes. A amabilidade e a competência mais elevadas também podem em conjunto ter eliciado a admiração à agente consoante o Stereotype Content Model (Cuddy et al., 2007; Fiske et al., 2002), que demonstra este nexos causal para agentes humanos, mas não conhecemos estudos que comprovem esta relação para agentes robóticos. Estas inferências quanto à admiração merecem investigações futuras.

#### **5.5. Agentes Artificiais, Robôs Humanoides e Cuidados de Saúde**

Consoante Reeves et al. (2020), os índices de amabilidade e de competência são homogêneos nas robôs de nossa escala de semelhança humana, o que nos permite descartar estas variáveis enquanto confundentes relativamente à dimensão da semelhança humana. As correlações mostraram que quanto maior a semelhança humana da robô imaginada, mais os participantes atribuíram competência, amabilidade e confiabilidade à agente robô.

Estes resultados são consistentes com o estudo de Phillips et al. (2018), que relatam que a aparência dos robôs talvez desempenhe um papel importante relativo aos julgamentos morais e pode influenciar a percepção das pessoas sobre seus demais atributos. Os robôs que se assemelham aos humanos parecem fornecer às pessoas uma sensação de familiaridade. A simulação de comportamentos (olhar, voz, expressões faciais, gestos) são importantes na interação social e na comunicação através do processo cognitivo do antropomorfismo (Giger et al., 2019; Devillers, 2021). O nosso estudo é também consistente com as evidências de Waytz et al. (2010) quando indicam que o antropomorfismo prediz a percepção de atributos como a confiabilidade e a capacidade num agente artificial. A capacidade em reconhecer a habilidade de outros seres humanos em entender as coisas e ter afetos e intenções (ToM) é essencial para as interações sociais, e o motivo da humanização dos robôs geralmente é ligado à preferência

humana pelas interações antropocêntricas (Giger et al., 2019). Os nossos resultados são também convergentes aos de Christoforakos et al. (2021), que mostraram um efeito positivo da competência e da amabilidade percebida nos robôs, moderado pelo antropomorfismo, na atribuição de confiabilidade aos robôs.

Quanto à semelhança humana atribuída pelos participantes à robô mais adequada à tarefa de enfermagem, o efeito significativo do tipo de agente pode ter sido originado pela dupla exposição dos participantes às fotos das robôs na condição de agente robótica, que resultou num maior índice de semelhança humana do que a condição de agente humana. Considerando que os participantes tinham escolhido uma determinada modelo de robô enquanto robô imaginada, a nova questão sobre a robô mais adequada à tarefa de enfermagem pode ter levado os participantes a optar por uma modelo de maior semelhança humana que a anteriormente escolhida.

Verificámos ainda que mais do que 50% de nossos participantes imaginaram robôs altamente semelhantes aos humanos como protagonistas dos dilemas morais descritos, e mais do que 75% escolheram robôs de elevado índice de semelhança humana como sendo os mais adequados à tarefa dos cuidados de saúde. Estes resultados são consistentes com os estudos de Waytz et al. (2010) quanto ao antropomorfismo e ao de Broadbent et al. (2013) que verificou a preferência por um robô que exibisse um rosto humano (em vez de um robô sem rosto humano) nas tarefas relativas à saúde, inferindo atributos positivos como a amabilidade.

Por contraste, estes resultados relativos à robô imaginada e à robô adequada à tarefa de enfermagem parecem contrariar a hipótese de Uncanny Valley (Mori, 1970), que sugere que a semelhança humana imperfeita de um robô pode evocar sentimentos de inquietação e estranheza em observadores humanos (Cheetham, 2017; Devillers, 2021). Apesar de as nossas evidências serem indiretas, considerando que os participantes escolheram os robôs mais próximos do que tinham imaginado, ou percebido como mais adequados à tarefa de enfermagem, a elevada percentagem de escolha de robôs mais semelhantes aos humanos é uma informação importante para o desenho de futuros robôs, mas merece também investigações e reflexões futuras.

A tecnologia e a humanidade são inexoravelmente mescladas na prática da enfermagem e dos cuidados de saúde. A tecnologia pode ser a mais básica possível (e.g., uma simples atadura), mas sempre fez parte do ofício do cuidar da saúde, que depende de competências e habilidades desenvolvidas através da experiência. Geralmente, o cuidado está associado ao calor humano, à amabilidade, ao ato muitas vezes intangível e subjetivo de nutrir, e a tecnologia, fria, racional, associada à competência e à confiabilidade, percebida como um instrumento concreto para



alcançar um fim (Mol et al., 2010). Entretanto, as tecnologias de saúde e os seus utilizadores (re)configuram uns aos outros consoante as correspondentes práticas de utilização, e muitas vezes as relações com estas tecnologias podem não ser apenas funcionais, mas também sociais e afetivas. Pols e Moser (2009), num estudo qualitativo sobre agentes robóticos (e.g., Aji, o cão robô; Health Buddy) utilizados em cuidados de saúde, analisaram as relações entre estas tecnologias e os seus utilizadores, e argumentaram não existir uma oposição entre a tecnologia “fria” e os cuidados “quentes”. Relatam que é possível a existência de relações de afeto entre os seres humanos e os agentes tecnológicos, de forma a atenuar distinções entre cuidados de saúde “frios” e “calorosos”, ludicidade e seriedade, tecnologias afetivas vs. racionais.

Os agentes artificiais têm potencialidades úteis no suporte aos agentes de saúde humanos, cujas competências e conhecimentos provavelmente não são passíveis de substituição por robôs sociais (Orians & Reisach, 2017). Os potenciais usuários advertem que os robôs de cuidados não devem substituir, mas complementar os serviços dos agentes humanos (Johansson-Pajala et al., 2020). Não obstante a competência e a confiabilidade dos robôs e da IA, há sobretudo barreiras legais e éticas, sendo um exemplo os possíveis laços afetivos estabelecidos por humanos vulneráveis com agentes artificiais (European Parliament, 2021). As tecnologias não funcionam sozinhas e dependem do trabalho humano, que não é apenas racional ou cognitivo, mas envolve uma praxis corporificada, imbuída de emoções. É presumível que os cuidados de saúde requeiram mais a atenção meticulosa e a flexibilidade do que a decisão firme e o julgamento imparcial: são um trabalho sobre a fragilidade do ser humano e pressupõem agir no sentido de melhorar a vida (Mol et al., 2010).

Os cuidados oferecidos por um agente humano tem duas dimensões, a relação de *mutualidade (caring about)* e a relação técnica-instrumental (*caring for*). Vários autores defendem que um agente robô não é capaz de prover a mutualidade humana, o que pode gerar uma dissonância e eliminar o significado do cuidado em si, sendo importante manter a transparência de que um robô social pode cuidar de um paciente, mas não pode se preocupar com ele (Mois & Beer, 2020; Parviainen & Pirhonen, 2017). O toque, o contato físico humano, é fundamental para os cuidados de saúde, em especial às pessoas idosas. Por outro lado, as tecnologias digitais como a robótica e a IA aplicadas à saúde não são necessariamente negativas, pois podem gerar ou reforçar as relações sociais entre os pacientes e os profissionais de saúde, dependendo da forma como são utilizadas (Breazeal, 2011; Mois & Beer, 2020; Parviainen & Pirhonen, 2017; Pols & Moser, 2009; Weber-Fiori et al., 2017).

## 5.6. Limitações e Estudos Futuros

O presente estudo apresenta algumas limitações, sendo uma delas a impossibilidade de utilizar as medidas do Moral Foundations Questionnaire para verificar possíveis interações com os julgamentos morais, pois em geral não foi possível obter consistência interna nas dimensões deste instrumento. Porém Cameron et al. (2015) observam que as correlações entre os diferentes fundamentos da Moral Foundations Questionnaire são mais elevadas do que a confiabilidade dentro de cada um dos fundamentos, o que pode explicar o porquê da reduzida consistência interna obtida. Os fundamentos morais parecem estar mais correlacionados uns com os outros do que com eles próprios, e o pré-requisito da distinção entre as dimensões pode não ter suporte empírico.

Aguinis e Bradley (2014) relatam uma suposta falta de realismo e a dificuldade de generalização dos resultados obtidos através das vinhetas experimentais ao sugerirem que os resultados usando vinhetas podem não se verificar fora da situação experimental. Bauman et al. (2014) advertem que as investigações por meio de dilemas sacrificiais envolvendo condições excessivamente simples e descontextualizadas podem apresentar baixo realismo experimental, psicológico e ecológico, e por estes motivos vir a mostrar baixa validade externa. Ao não representar as questões éticas encontradas no mundo real, tais dilemas podem não envolver os mesmos processos psicológicos que as situações morais cotidianas. No entanto, mesmo considerando estes pressupostos, destacamos que o nosso dilema moral apresenta um cenário relativo a cuidados de saúde que pode ser encontrado no dia-a-dia das pessoas, de forma a apresentar uma melhor validade externa. Com a intenção de controlar esta potencial limitação, avaliamos o grau de compreensão e de credibilidade das vinhetas apresentadas. Os resultados mostraram índices positivos relatados pelos participantes, sendo que 81.9% indicaram ter compreendido bem a vinheta e 84.1% concordou sobre o realismo da vinheta. Além disto, a nossa questão de validação eliminou os participantes aos quais não havia ficado clara a situação na qual decorreu o dilema (um ambiente domiciliar), de forma a assegurar o real entendimento da situação. Asseguramos assim a compreensão das nossas vinhetas, que simularam cenários percebidos como realistas e semelhantes a situações conhecidas pelos participantes. A manipulação realizada por meio das oito diferentes vinhetas eliciou efeitos observáveis através dos resultados alcançados (validade interna), efeitos que hipoteticamente poderão ter paralelo na realidade e serem passíveis de generalização (validade externa) às condições do mundo real (Evans et al., 2015). A opção pelo recurso a vinhetas experimentais no nosso estudo foi ainda justificada pelos seguintes motivos: foi útil para manipular variáveis independentes, permitindo analisar verificar relações de causalidade; possibilitou a inclusão de fatores relevantes para a

investigação e o controlo de variáveis confundentes; e facilitou a criação de cenários hipotéticos na abordagem de uma temática sensível (Aguinis & Bradley, 2014; Evans et al., 2015). Especialmente no contexto dos cuidados de saúde, muitas vezes é impossível ou pouco ético manipular experimentalmente a tomada de decisão e as ações dos agentes de saúde em relação aos pacientes, e as vinhetas são um método adequado para avaliar aspetos como a qualidade dos cuidados prestados e a conformidade com os princípios éticos (Evans et al., 2015). Não obstante, seria interessante o recurso a desenhos experimentais que incluíssem robôs físicos em cenários semelhantes, mas num contexto real de cuidados de saúde, salvaguardando toda a conformidade com os princípios éticos.

Outra limitação de nossas vinhetas é a de que a agente de saúde ofereceu uma explicação (argumentação sobre o benefício ou prejuízo à saúde) do porquê não respeitar a autonomia na condição de beneficência/não-maleficência, mas não explicou a razão de acatar a decisão da paciente na condição do respeito à autonomia. Este detalhe pode ter originado uma assimetria em decorrência de uma explicação ser oferecida numa condição, mas não em outra, o que pode ter influenciado a avaliação dos participantes em relação aos atributos das agentes (Martin & Cushman, 2016; Reeder et al., 2002). Investigações futuras poderiam utilizar vinhetas nas quais são apresentadas (ou não) explicações consistentes para ambas as decisões éticas. Desta forma, seria possível elucidar se há uma diferença mensurável nos julgamentos morais no caso da existência ou não de uma justificação para uma ação.

A nossa investigação procurou controlar o género, tendo sido restrita ao género feminino no par agente-paciente. Estudos futuros devem ser realizados para verificar a possível influência do género de agentes de saúde e pacientes em dilemas morais envolvendo relacionamentos na área dos cuidados de saúde mediados por agentes artificiais. A meta-análise de Hall & Roter (2002) mostra claras diferenças entre géneros nas relações agente-paciente, e Johanson et al. (2020) trazem evidências de que as variações de género dos agentes (humanos e robôs) e dos pacientes influenciam os resultados da interação entre eles.

O presente estudo não abordou o princípio ético da justiça, o que pode ser visto como uma limitação. A revisão sistemática de Murphy et al. (2021) refere a ausência de literatura científica sobre a ética da IA na saúde global, especialmente nos países de baixo e médio rendimento, e apontam a necessidade de investigação nessa área a fim de um desenvolvimento e uma implementação da IA na saúde pública e global de forma socialmente equitativa.

Outra sugestão para futuras investigações é a utilização de um desenho intra-sujeitos, dado que no presente estudo se optou por um desenho inter-sujeitos. Uma possibilidade seria variar a ordem de apresentação do agente (e.g., enquanto um grupo de participantes analisa uma

vinheta com um agente robótico e logo após uma vinheta com um agente humano, outro grupo seria exposto às vinhetas na ordem inversa). Desta forma, será possível a comparação entre os participantes que tenham se confrontado com ambos os tipos de agente.

Finalmente, seria interessante estender o presente estudo para outras culturas, países e linguagens, considerando a limitação decorrente da utilização da língua portuguesa, através da qual envolvemos as culturas do Brasil e de Portugal.

## 6. Conclusão

A presente investigação procurou investigar os julgamentos morais acerca de agentes de saúde humanos e robôs sociais responsáveis por realizar tarefas de cuidados de saúde, incluindo tomadas de decisão éticas e argumentações sobre os efeitos do uso de uma medicação.

No geral, os resultados mostraram algumas diferenças nos julgamentos efetuados aos dois tipos de agente. Verificou-se que os julgamentos de aceitabilidade moral foram semelhantes para as duas agentes, a humana e a robô. No entanto, a responsabilidade moral imputada à agente humana foi superior à atribuída ao robô, de forma análoga ao estudo de Kahn et al. (2012), realizado através da interação dos participantes com um robô real, e ao de Laakasuo et al. (2019) que também usou da metodologia das vinhetas experimentais. Na avaliação dos atributos, a agente humana foi avaliada como mais amável do que a robô, mas em termos da avaliação de competência e de confiabilidade as percepções foram semelhantes. Estes resultados evidenciam a tendência para a atribuição de “personalidade” aos robôs mediante critérios similares aos que usam para seres humanos, mas de uma forma um pouco diferente para os robôs, evidência consistente com os estudos anteriores (de Graaf & Ben Allouch, 2016; de Graaf & Malle, 2019; Gamez et al., 2020; Kahn et al., 2012).

Destacou-se, no presente estudo, a importância da decisão ética para a maioria dos julgamentos e avaliações efetuadas, em detrimento do tipo de agente ou da argumentação. Este resultado relativo aos julgamentos de caráter moral também está presente nas evidências de Malle et al. (2016). Assim, independente do tipo de agente, as que respeitaram a autonomia da paciente foram julgadas com maior aceitabilidade moral por comparação às que optaram pela beneficência/não-maleficência. A nossa investigação, assim como a de Laakasuo et al. (2019), revelou uma preferência dos participantes por agentes (humanas ou robôs) que respeitam a autonomia da paciente.

Os nossos resultados sugerem uma clivagem entre a dimensão emocional da amabilidade e os aspetos predominantemente cognitivos da competência e da confiabilidade relativamente à decisão ética. Houve uma atribuição elevada de amabilidade às agentes que respeitaram a autonomia da paciente, porém com avaliações um pouco menores de competência e confiabilidade, de acordo com a dimensão do efeito estimado nestas análises. Em contrapartida, a opção pela beneficência/não-maleficência ao paciente influenciou a percepção de agentes um pouco mais confiáveis e competentes, mas menos amáveis. Salientamos assim que o efeito nas avaliações da amabilidade apresentou uma dimensão de efeito principal mais elevada e uma diferença maior entre as médias que os demais atributos no que se refere à influência da decisão ética, parecendo ser uma dimensão com resultados mais sólidos. Este resultado é consistente

com Melchert (2020), que defende a importância crucial dos atributos associados à amabilidade nos cuidados de saúde, às vezes maior do que a importância dos atributos técnicos. Embora a percepção da amabilidade pareça ser prioritária na avaliação das pessoas acerca dos potenciais riscos e benefícios na interação com um agente, o atributo da competência também é fundamental nos cuidados de saúde, ao indicar o nível de benefício que um agente pode oferecer ao paciente. Estes atributos podem ter importância variável consoante diferentes problemas: a amabilidade poderá ser mais relevante quando se trata de algum problema crónico e que precisa de tratamento ao longo do tempo (situação na qual podemos enquadrar nosso dilema moral), enquanto a competência parece primordial nos casos ambulatoriais agudos (Howe et al., 2019).

Com a finalidade de ilustrar a clivagem da decisão ética e dos atributos da agente – de um lado, o respeito à autonomia e a amabilidade, e do outro lado, a priorização da beneficência/não-maleficência e a competência e confiabilidade – resumimos um caso relatado por Mol (2008). Certa vez, um eticista lançou um dilema para uma equipa de saúde:

“Imagine um paciente de uma ala aberta num hospital psiquiátrico a não querer levantar-se da cama. Enquanto agente de saúde, permitiria que ele ficasse na cama ou não?”. Um psicoterapeuta aposentado respondeu que enviaria uma enfermeira para questionar o porquê de o paciente não levantar a fim de perceber o porquê desse comportamento, para dar-lhe atenção e deixá-lo falar, pois alguém que não quer sair da cama precisa de cuidados. Seria negligência deixá-lo na cama ou forçá-lo a levantar-se.

Há uma alternativa à lógica da polarização: a *ética prática dos cuidados* (Molterer et al., 2020). O nosso dilema ético revela uma clivagem e polarização e permite delinear os cuidados de saúde de outra forma. Uma paciente que se recusa a tomar um medicamento, fá-lo por alguma razão. Ao negar esta recusa, a agente invalida a vontade da paciente e anula a paciente enquanto indivíduo, o que pode ser interpretado como um comportamento pouco ou nada amável. Assim, a agente de saúde parece cumprir de modo competente e confiável a sua tarefa numa *lógica profissional dos cuidados*, mas a comunicação agente-paciente e o papel da agente numa *lógica relacional dos cuidados* são deixados de lado (Molterer et al., 2020). A não-aquiescência perante às preferências de um paciente, mesmo na intenção do benefício à saúde, é considerada uma forma de paternalismo (Beauchamp & Childress, 2019). Em contrapartida, o modelo consumerista de simplesmente respeitar a autonomia da paciente não parece totalmente adequado: apesar desta conduta da agente aparentar maior amabilidade, enfraquece tecnicamente seu papel profissional e sua competência e confiabilidade podem ser percebidas como inferiores.

O nosso dilema moral pode ser explicado pela díade utilitarismo-deontologia. Porém, nenhuma teoria ética pode ser considerada inferior ou derivada de outra, mas articulam visões diferentes, às vezes contraditórias, que devem ser consideradas nos cuidados de saúde. O estudo de agentes morais com foco apenas nesta díade pode deixar de lado conceitos relevantes de outras teorias éticas (Beauchamp & Childress, 2019; Zoshak & Dew, 2021). À medida que os riscos à saúde de um paciente aumentam devido à não adesão, é menos provável que exista a obrigação de respeitar a sua autonomia, e muitas vezes são o dever ou o caráter moral do agente (teoria da virtude) que estão em questão (Beauchamp & Childress, 2019). Cabe à agente de saúde acolher a não adesão da paciente e compreender o que a paciente quer comunicar com seu comportamento, interagindo através da lógica dos cuidados, numa ética prática dos cuidados (Koggel & Orme, 2010; McDermott, 2011; Mol, 2008; Mol et al., 2010; Molterer et al., 2020; Pettersen, 2011). O estabelecimento de uma relação de mutualidade é talvez a melhor solução para a construção conjunta de um procedimento que favoreça a saúde da paciente numa perspectiva biopsicossocial (Hall & Roter, 2011), sendo crucial que as idiosincrasias de um paciente possam ser mescladas com as preocupações técnicas do agente de saúde.





## 7. Referências

- Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods, 17*(4), 351–371. <https://doi.org/10.1177/1094428114547952>
- Alaieri, F., & Vellino, A. (2016). Ethical decision making in robots: Autonomy trust and responsibility. In A. Agah, J.-J. Cabibihan, A. M. Howard, M. A. Salichs, & H. He (Eds.), *Social Robotics* (Vol. 9979, pp. 159–168). Springer International Publishing. [https://doi.org/10.1007/978-3-319-47437-3\\_16](https://doi.org/10.1007/978-3-319-47437-3_16)
- American Nurses Association. (2021, March). *What is Nursing & What do nurses do?* <https://www.nursingworld.org/practice-policy/workforce/what-is-nursing/>
- Ancis, J. R. (2020). The age of cyberpsychology: An overview. *Technology, Mind, and Behavior, 1*(1), 1–15. <https://doi.org/10.1037/tmb0000009>
- Anderson, M., Anderson, S. L., & Armen, C. (2006, July 16–20). *MedEthEx: A prototype medical ethics advisor*. [Paper presentation] Proceedings of the 18th Conference on Innovative Applications of Artificial Intelligence, Boston, Massachusetts, United States. <https://www.aaai.org/Library/AAAI/2006/aaai06-292.php>
- Anderson, M. & Anderson, S. L. (2018). GenEth: A general ethical dilemma analyzer. *Paladyn, Journal of Behavioral Robotics, 9*(1), 337–357. <https://doi.org/10.1515/pjbr-2018-0024>
- Ashcroft, R. E. (2012). Health Technology Assessment. In R. F. Chadwick (Ed.), *Encyclopedia of Applied Ethics* (2nd ed., vol. 2, pp. 556–565). Academic Press.
- Baird, C. A. (2012). Everyday Ethics. In R. F. Chadwick (Ed.), *Encyclopedia of Applied Ethics* (2nd ed., vol. 2, pp. 208–216). Academic Press.
- Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass, 8*(9), 536–554. <https://doi.org/10.1111/spc3.12131>
- Beauchamp, T. L., & Childress, J. F. (2019). *Principles of Biomedical Ethics*. (8th ed.). Oxford University Press.
- Benyamini, Y. (2011). Health and illness perceptions. In H. S. Friedman (Ed.), *The Oxford Handbook of Health Psychology* (pp. 281–314). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195342819.001.0001>
- Breazeal, C. (2011). Social robots for health applications. *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 5368–5371*. <https://doi.org/10.1109/IEMBS.2011.6091328>
- Broadbent, E., Kumar, V., Li, X., Sollers, J., Stafford, R. Q., MacDonald, B. A., & Wegner, D. M. (2013). Robots with display screens: A robot with a more humanlike face display is perceived to have more mind and a better personality. *PLoS ONE, 8*(8), e72589. <https://doi.org/10.1371/journal.pone.0072589>
- Cameron, C. D., Lindquist, K. A., & Gray, K. (2015). A constructionist review of morality and emotions: No evidence for specific links between moral content and discrete emotions. *Personality and Social Psychology Review, 19*(4), 371–394. <https://doi.org/10.1177/1088868314566683>
- Carruthers, P. & Smith, P. K. (Eds.). (1998). *Theories of theories of mind* (Reprinted). Cambridge University Press.
- Cheetham, M. (2017). The Uncanny Valley Hypothesis and beyond [Editorial]. *Frontiers in Psychology, 8*, 1738. <https://doi.org/10.3389/fpsyg.2017.01738>
- Christoforakos, L., Gallucci, A., Surmava-Große, T., Ullrich, D., & Diefenbach, S. (2021). Can robots earn our trust the same way humans do? A systematic exploration of competence, warmth, and anthropomorphism as determinants of trust development in HRI. *Frontiers in Robotics and AI, 8*. <https://doi.org/10.3389/frobt.2021.640444>

- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, 92(4), 631–648. <https://doi.org/10.1037/0022-3514.92.4.631>
- Dai, Z., & MacDorman, K. F. (2018). The doctor’s digital double: How warmth, competence, and animation promote adherence intention. *PeerJ Computer Science*, 4, e168. <https://doi.org/10.7717/peerj-cs.168>
- Damásio, A. (2019). *O Erro de Descartes* [Descartes’ Error] (6th ed.). Temas e Debates.
- de Graaf, M. M. A., & Ben Allouch, S. (2016). Anticipating our future robot society: The evaluation of future robot applications from a user’s perspective. *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 755–762. <https://doi.org/10.1109/ROMAN.2016.7745204>
- de Graaf, M. M. A., & Malle, B. F. (2019). People’s explanations of robot behavior subtly reveal mental state inferences. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 239–248. <https://doi.org/10.1109/HRI.2019.8673308>
- Devillers, L. (2021). Human–robot interactions and affective computing: The ethical implications. In J. von Braun, M. S. Archer, G. M. Reichberg, & M. Sánchez Sorondo (Eds.), *Robotics, AI, and Humanity* (pp. 205–211). Springer International Publishing. [https://doi.org/10.1007/978-3-030-54173-6\\_17](https://doi.org/10.1007/978-3-030-54173-6_17)
- Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science*, 40(3), 434–449. <https://doi.org/10.1007/s11747-011-0300-3>
- Dodig-Crnkovic, G., & Çürüklü, B. (2012). Robots: Ethical by design. *Ethics and Information Technology*, 14(1), 61–71. <https://doi.org/10.1007/s10676-011-9278-2>
- Dorison, C., Lerner, J. S., Heller, B. H., Rothman, A., Kawachi, I. I., Wang, K., ... Coles, N. A. (2020, April 16). A global test of message framing on behavioural intentions, policy support, information seeking, and experienced anxiety during the COVID-19 pandemic. <https://doi.org/10.31234/osf.io/sev kf>
- Duke, C. C., Lynch, W. D., Smith, B., & Winstanley, J. (2015). Validity of a new patient engagement measure: The Altarum Consumer Engagement (ACE) Measure™. *The Patient-Patient-Centered Outcomes Research*, 8(6), 559–568. <https://doi.org/10.1007/s40271-015-0131-2>
- Engel, G. L. (2012). The need for a new medical model: A challenge for biomedicine. *Psychodynamic psychiatry*, 40(3), 377-396. <https://doi.org/10.1521/pdps.2012.40.3.377>
- Eshleman, A. (2014). Moral Responsibility. In R. S. Dillon, & [Respect] E. N. Zalta (Eds.) *The Stanford Encyclopedia of Philosophy*. [https://pilot scholars.up.edu/phl\\_facpubs/1](https://pilot scholars.up.edu/phl_facpubs/1)
- European Parliament (2021). *Report with recommendations to the Commission on Civil Law Rules on Robotics* (2015/2103/INL). Retrieved 2 July 2021, from [http://www.europarl.europa.eu/doceo/document/A-8-2017-0005\\_EN.html](http://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html)
- Evans, S. C., Roberts, M. C., Keeley, J. W., Blossom, J. B., Amaro, C. M., Garcia, A. M., Stough, C. O., Canter, K. S., Robles, R., & Reed, G. M. (2015). Vignette methodologies for studying clinicians’ decision-making: Validity, utility, and application in ICD-11 field studies. *International Journal of Clinical and Health Psychology*, 15(2), 160–170. <https://doi.org/10.1016/j.ijchp.2014.12.001>
- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics*. (5th ed.). Sage.
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and

- competition. *Journal of personality and social psychology*, 82(6), 878. <https://doi.org/10.1037/0022-3514.82.6.878>
- Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3–4), 143–166. [https://doi.org/10.1016/S0921-8890\(02\)00372-X](https://doi.org/10.1016/S0921-8890(02)00372-X)
- Gallagher, K. M., & Updegraff, J. A. (2011). Health message framing effects on attitudes, intentions, and behavior: A meta-analytic review. *Annals of Behavioral Medicine*, 43(1), 101–116. <https://doi.org/10.1007/s12160-011-9308-7>
- Gamez, P., Shank, D. B., Arnold, C., & North, M. (2020). Artificial virtue: The machine question and perceptions of moral character in artificial moral agents. *AI & SOCIETY*, 35(4), 795–809. <https://doi.org/10.1007/s00146-020-00977-1>
- Giger, J., Piçarra, N., Alves-Oliveira, P., Oliveira, R., & Arriaga, P. (2019). Humanization of robots: Is it really such a good idea? *Human Behavior and Emerging Technologies*, 1(2), 111–123. <https://doi.org/10.1002/hbe2.147>
- Gillon, R. (1994). Medical ethics: Four principles plus attention to scope. *BMJ*, 309(6948), 184–188. <https://doi.org/10.1136/bmj.309.6948.184>
- Gleichgerricht, E., & Young, L. (2013). Low levels of empathic concern predict utilitarian moral judgment. *PLoS ONE*, 8(4), e60418. <https://doi.org/10.1371/journal.pone.0060418>
- Graham, J., Nosek, B., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, 101, 366–385. <https://doi.org/10.1037/a0021847>
- Guimarães, R., & Cabral, J. A. S. (1997). *Estatística* [Statistics]. McGraw Hill.
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Pantheon.
- Hall, J. A., & Roter, D. L. (2002). Do patients talk differently to male and female physicians? *Patient Education and Counseling*, 48(3), 217–224. [https://doi.org/10.1016/s0738-3991\(02\)00174-x](https://doi.org/10.1016/s0738-3991(02)00174-x)
- Hall, J. A., & Roter, D. L. (2011). Physician–patient communication. In H. S. Friedman (Ed.), *The Oxford Handbook of Health Psychology* (pp. 317–346). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195342819.001.0001>
- Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede Model in context. *Online Readings in Psychology and Culture*, 2(1). <https://doi.org/10.9707/2307-0919.1014>
- Howe, L. C., Leibowitz, K. A., & Crum, A. J. (2019). When your doctor “Gets it” and “Gets you”: The critical role of competence and warmth in the patient-provider interaction. *Frontiers in Psychiatry*, 10, 475. <https://doi.org/10.3389/fpsy.2019.00475>
- Järvik, M. (2003). How to understand moral responsibility? *Trames: A Journal of the Humanities and Social Sciences*, 7(3), 147–163.
- Johanson, D. L., Ahn, H. S., & Broadbent, E. (2020). Improving interactions with healthcare robots: A review of communication behaviours in social and healthcare contexts. *International Journal of Social Robotics*. <https://doi.org/10.1007/s12369-020-00719-9>
- Johansson-Pajala, R. M., Thommes, K., Hoppe, J. A., Tuisku, O., Hennala, L., Pekkarinen, S., ... & Gustafsson, C. (2020). Care robot orientation: what, who and how? Potential users’ perceptions. *International Journal of Social Robotics*, 1–15. <https://doi.org/10.1007/s12369-020-00619-y>
- Kahn, P. H., Reichert, A. L., Gary, H. E., Kanda, T., Ishiguro, H., Shen, S., Ruckert, J. H., & Gill, B. (2011). The new ontological category hypothesis in human-robot interaction. *Proceedings of the 6th International Conference on Human-Robot Interaction - HRI '11*, 159. <https://doi.org/10.1145/1957656.1957710>
- Kahn, P. H., Severson, R. L., Kanda, T., Ishiguro, H., Gill, B. T., Ruckert, J. H., Shen, S., Gary, H. E., Reichert, A. L., & Freier, N. G. (2012). Do people hold a humanoid robot morally

- accountable for the harm it causes? *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI '12*, 33. <https://doi.org/10.1145/2157689.2157696>
- Kahn, P. H., Gary, H. E., & Shen, S. (2013). Children's social relationships with current and near-future robots. *Child Development Perspectives*, 7(1), 32–37. <https://doi.org/10.1111/cdep.12011>
- Kaplan, R. M. (2011). Uncertainty, variability, and resource allocation in the health care decision process. In H. S. Friedman (Ed.), *The Oxford Handbook of Health Psychology* (pp. 95–118). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195342819.001.0001>
- Khan, Z. H., Siddique, A., & Lee, C. W. (2020). Robotics utilization for healthcare digitization in global COVID-19 management. *International Journal of Environmental Research and Public Health*, 17(11), 3819. <https://doi.org/10.3390/ijerph17113819>
- Kim, M., Yi, S., & Lee, D. (2019). Between living and nonliving: Young children's animacy judgments and reasoning about humanoid robots. *PLOS ONE*, 14(6), 1–19. <https://doi.org/10.1371/journal.pone.0216869>
- Kim, M., Park, B., & Young, L. (2020). The psychology of motivated versus rational impression updating. *Trends in Cognitive Sciences*, 24(2), 101–111. <https://doi.org/10.1016/j.tics.2019.12.001>
- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. (4th ed.). Guilford Press.
- Koggel, C., & Orme, J. (2010). Care ethics: New theories and applications. *Ethics and Social Welfare*, 4(2), 109–114. <https://doi.org/10.1080/17496535.2010.484255>
- Laakasuo, M., Kunnari, A., Palomäki, J., Rauhalala, S., Koverola, M., Lehtonen, N., Halonen, J., Repo, M., Visala, A., & Drosinou, M. (2019). *Moral psychology of nursing robots – humans dislike violations of patient autonomy but like robots disobeying orders* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/bkhyq>
- Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual review of psychology*, 66. <https://doi.org/10.1146/annurev-psych-010213-115043>
- Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences*, 8(12), 528–533. <https://doi.org/10.1016/j.tics.2004.10.001>
- Locsin, R. C., & Ito, H. (2018). Can humanoid nurse robots replace human nurses? *Journal of Nursing*, 5(1):1. <https://doi.org/10.7243/2056-9157-5-1>
- Loiacono, E. T., Watson, R. T., & Goodhue, D. L. (2002). WebQual: A measure of website quality. *Marketing theory and applications*, 13(3), 432–438.
- Malle, B. F. (2021). Moral judgments. *Annual Review of Psychology*, 72(1), 293–318. <https://doi.org/10.1146/annurev-psych-072220-104358>
- Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., & Cusimano, C. (2015). Sacrifice one for the good of many? In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 117–124. <https://doi.org/10.1145/2696454.2696458>
- Malle, B. F., Scheutz, M., Forlizzi, J., & Voiklis, J. (2016). Which robot am I thinking about? The impact of action and appearance on people's evaluations of a moral robot. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 125–132. <https://doi.org/10.1109/HRI.2016.7451743>
- Malle, B. F., Magar, S. T., & Scheutz, M. (2019). AI in the sky: How people morally evaluate human and machine decisions in a lethal strike dilemma. In M. I. Aldinhas Ferreira, J. Silva Sequeira, G. Singh Virk, M. O. Tokhi, & E. E. Kadar (Eds.), *Robotics and Well-Being* (Vol. 95, pp. 111–133). Springer International Publishing. [https://doi.org/10.1007/978-3-030-12524-0\\_11](https://doi.org/10.1007/978-3-030-12524-0_11)

- Malle, B. F., & Ullman, D. (2021). A multidimensional conception and measure of human-robot trust. In C. S. Nam & J. B. Lyons (Eds.), *Trust in Human-Robot Interaction* (pp. 3–25). Academic Press. <https://doi.org/10.1016/B978-0-12-819472-0.00001-0>
- Martin, J. W., & Cushman, F. (2016). Why we forgive what can't be controlled. *Cognition*, *147*, 133–143. <https://doi.org/10.1016/j.cognition.2015.11.008>
- McDermott, S. (2011). Ethical decision making in situations of self-neglect and squalor among older people. *Ethics and Social Welfare*, *5*(1), 52–71. <https://doi.org/10.1080/17496535.2011.546179>
- Melchert, T. P. (2020). *Foundations of health service psychology: An evidence-based biopsychosocial approach*. (2nd ed.). Academic Press.
- Mois, G., & Beer, J. M. (2020). Robotics to support aging in place. In *Living with Robots* (pp. 49–74). Elsevier. <https://doi.org/10.1016/B978-0-12-815367-3.00003-7>
- Mol, A. (2008). *The logic of care: Health and the problem of patient choice*. Routledge.
- Mol, A., Moser, I., & Pols, J. (2010). Care: Putting practice into theory. In A. Mol, I. Moser, & J. Pols (Eds.), *Care in Practice* (pp. 7–26). transcript Verlag. <https://doi.org/10.14361/transcript.9783839414477.7>
- Molterer, K., Hoyer, P., & Steyaert, C. (2020). A practical ethics of care: Tinkering with different 'goods' in residential nursing homes. *Journal of Business Ethics*, *165*(1), 95–111. <https://doi.org/10.1007/s10551-018-04099-z>
- Mori, M. (1970) The Uncanny Valley. *Energy*, *7*:33–35
- Murphy, K., Di Ruggiero, E., Upshur, R., Willison, D. J., Malhotra, N., Cai, J. C., ... & Gibson, J. (2021). Artificial intelligence for good health: A scoping review of the ethics literature. *BMC medical ethics*, *22*(1), 1–17. <https://doi.org/10.1186/s12910-021-00577-8>
- Nabi, R. L., Walter, N., Oshidary, N., Endacott, C. G., Love-Nichols, J., Lew, Z. J., & Aune, A. (2020). Can emotions capture the elusive gain-loss framing effect? A meta-analysis. *Communication Research*, *47*(8), 1107–1130. <https://doi.org/10.1177/0093650219861256>
- O'Keefe, D. J., & Jensen, J. D. (2007). The relative persuasiveness of gain-framed loss-framed messages for encouraging disease prevention behaviors: A meta-analytic review. *Journal of health communication*, *12*(7), 623–644. <https://doi.org/10.1080/10810730701615198>
- Orians, W., & Reisach, U. (2017). Wissenstransfer in der Kranken- und Altenpflege: Möglichkeiten und Grenzen der Digitalisierung von Wissen [Knowledge Transfer in Nursing and Care for the Elderly: Possibilities and Limits of Digitizing of Knowledge]. In M. A. Pfannstiel, S. Krammer, & W. Swoboda (Eds.), *Digitale Transformation von Dienstleistungen im Gesundheitswesen III* (pp. 33–54). Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-13642-0\\_3](https://doi.org/10.1007/978-3-658-13642-0_3)
- Parviainen, J., & Pirhonen, J. (2017). Vulnerable bodies in human-robot interaction: Embodiment as ethical issue in robot care for the elderly. *Transformations*, *29* (special issue: 'Social robots: Human-machine configurations'), 104–115.
- Peck, M. L. (1992). The future of nursing in a technological age: Computers, robots, and TLC. *Journal of Holistic Nursing*, *10*(2), 183–191. <https://doi.org/10.1177/089801019201000208>
- Pepito, J. A., & Locsin, R. (2019). Can nurses remain relevant in a technologically advanced future? *International journal of nursing sciences*, *6*(1), 106–110. <https://doi.org/10.1016/j.ijnss.2018.09.013>
- Pettersen, T. (2011). The ethics of care: Normative structures and empirical implications. *Health Care Analysis*, *19*(1), 51–64. <https://doi.org/10.1007/s10728-010-0163-7>
- Phillips, E., Zhao, X., Ullman, D., & Malle, B. F. (2018). What is human-like? Decomposing robot human-like appearance using the Anthropomorphic roBOT (ABOT) Database. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 105–113. <https://doi.org/10.1145/3171221.3171268>

- Piçarra, N. (2015). *Predicting intention to work with social robots* [Doctoral dissertation, Universidade do Algarve]. Repositório da Universidade do Algarve. <http://sapientia.ualg.pt/handle/10400.1/6854>
- Piçarra, N., Giger, J.-C., Pochwatko, G., & Gonçalves, G. (2015). Validation of the Portuguese version of the Negative Attitudes towards Robots Scale. *European Review of Applied Psychology*, 65(2), 93–104. <https://doi.org/10.1016/j.erap.2014.11.002>
- Pols, J., & Moser, I. (2009). Cold technologies versus warm care? On affective and social relations with and through care technologies. *Alter, European Journal of Disability Research*, 3(2), 159–178. <https://doi.org/10.1016/j.alter.2009.01.003>
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515–526. Cambridge Core. <https://doi.org/10.1017/S0140525X00076512>
- Pritchard, J. (2012). Codes of ethics. In R. F. Chadwick (Ed.), *Encyclopedia of Applied Ethics* (2nd ed., vol. 1, pp. 494–499). Academic Press.
- Rasche, C. (2017). Digitaler Gesundheitswettbewerb: Strategien, Geschäftsmodelle, Kompetenzanforderungen [Digital health competition: Strategies, business models, competence requirements]. In M. A. Pfannstiel, P. Da-Cruz, & H. Mehlich (Eds.), *Digitale Transformation von Dienstleistungen im Gesundheitswesen I* (pp. 1–30). Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-12258-4\\_1](https://doi.org/10.1007/978-3-658-12258-4_1)
- Reeder, G. D., Kumar, S., Hesson-McInnis, M. S., & Trafimow, D. (2002). Inferences about the morality of an aggressor: The role of perceived motive. *Journal of Personality and Social Psychology*, 83(4), 789–803. <https://doi.org/10.1037/0022-3514.83.4.789>
- Reeves, B., Hancock, J., & Liu, X. (2020). Social robots are like real people: First impressions, attributes, and stereotyping of social robots. *Technology, Mind, and Behavior*, 1(1). <https://doi.org/10.1037/tmb0000018>
- Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach*. (4th ed.). Pearson.
- Sasson, S. (2000). Beneficence versus respect for autonomy. *Journal of Gerontological Social Work*, 33(1), 5–16. [https://doi.org/10.1300/J083v33n01\\_02](https://doi.org/10.1300/J083v33n01_02)
- Sauer, H. (2017). Between facts and norms: Ethics and empirical moral psychology. In B. G. Voyer & T. Tarantola (Eds.), *Moral Psychology* (pp. 5–27). Springer International Publishing. [https://doi.org/10.1007/978-3-319-61849-4\\_2](https://doi.org/10.1007/978-3-319-61849-4_2)
- Scheuer, E. (2017). Wie Medical-Decision-Support-Systeme die Arzt-Patient-Beziehung verändern – Digitalisierung von Informationen führt zu einer erhöhten Autonomie des Patienten [How medical decision support systems are changing the doctor-patient relationship – digitization of information leads to increased patient autonomy]. In M. A. Pfannstiel, P. Da-Cruz, & H. Mehlich (Eds.), *Digitale Transformation von Dienstleistungen im Gesundheitswesen I* (pp. 311–321). Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-12258-4\\_20](https://doi.org/10.1007/978-3-658-12258-4_20)
- SDG Good Practices | Department of Economic and Social Affairs. (n.d.). Retrieved 6 July 2021, from <https://sdgs.un.org/sdg-good-practices>
- Sharkey, A. (2017). Can robots be responsible moral agents? And why should we care? *Connection Science*, 29(3), 210–216. <https://doi.org/10.1080/09540091.2017.1313815>
- Sommer, K., Nielsen, M., Draheim, M., Redshaw, J., Vanman, E. J., & Wilks, M. (2019). Children’s perceptions of the moral worth of live agents, robots, and inanimate objects. *Journal of Experimental Child Psychology*, 187, 104656. <https://doi.org/10.1016/j.jecp.2019.06.009>
- Taylor, S. E. (2011). Social support: A review. In H. S. Friedman (Ed.), *The Oxford Handbook of Health Psychology* (pp. 189–214). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195342819.001.0001>
- Triskiel, J. (2016). Psychology instead of ethics? Why psychological research is important but

- cannot replace ethics. In C. Brand (Ed.), *Dual-Process Theories in Moral Psychology* (pp. 77–98). Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-12053-5\\_4](https://doi.org/10.1007/978-3-658-12053-5_4)
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453–458. <https://doi.org/10.1126/science.7455683>
- Ullman, D., & Malle, B. F. (2018). What does it mean to trust a robot? Steps toward a multidimensional measure of trust. *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 263–264. <https://doi.org/10.1145/3173386.3176991>
- Vandemeulebroucke, T., de Casterle, B. D., & Gastmans, C. (2018). How do older adults experience and perceive socially assistive robots in aged care: A systematic review of qualitative evidence. *Aging & mental health*, 22(2), 149–167. <https://doi.org/10.1080/13607863.2017.1286455>
- VandenBos, G. R. (Ed.). (2015). *APA dictionary of psychology* (2nd ed.). American Psychological Association. <https://doi.org/10.1037/14646-000>
- Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., ... & Nerini, F. F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature communications*, 11(1), 1–10. <https://doi.org/10.1038/s41467-019-14108-y>
- Voiklis, J., Kim, B., Cusimano, C., & Malle, B. F. (2016). Moral judgments of human vs. robot agents. *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 775–780. <https://doi.org/10.1109/ROMAN.2016.7745207>
- Waytz, A., Cacioppo, J., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science*, 5(3), 219–232. <https://doi.org/10.1177/1745691610369336>
- Weber-Fiori, B., Stähle, B., Pfiffner, S., Reiner, B., Ertel, W., & Winter, M. H.-J. (2017). Marvin, ein Assistenzroboter für Menschen mit körperlicher Behinderung im praktischen Einsatz [Marvin, an assistance robot for people with physical disabilities in practical use]. In M. A. Pfannstiel, S. Krammer, & W. Swoboda (Eds.), *Digitale Transformation von Dienstleistungen im Gesundheitswesen III* (pp. 269–285). Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-13642-0\\_17](https://doi.org/10.1007/978-3-658-13642-0_17)
- Williams, G. (2012). Responsibility. In R. F. Chadwick (Ed.), *Encyclopedia of Applied Ethics* (2nd ed., vol. 3, pp. 821–828). Academic Press.
- Yu, K.-H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature Biomedical Engineering*, 2(10), 719–731. <https://doi.org/10.1038/s41551-018-0305-z>
- Zickfeld, J., van de Ven, N., Pich, O., Schubert, T. W., Berkessel, J., Pizarro Carrasco, J. J., ... Orjiakor, C. (2020, November 8). Tears trigger the intention to offer social support: A systematic investigation of the interpersonal effects of emotional crying across 41 countries. <https://doi.org/10.31234/osf.io/p7s5v>
- Zoshak, J., & Dew, K. (2021). Beyond Kant and Bentham: How ethical theories are being used in artificial moral agents. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15. <https://doi.org/10.1145/3411764.3445102>





## 8. Anexo A. Questionário

---

Início do bloqueio: Consentimento Informado



### Tecnologia e Cuidados de Saúde

A presente investigação tem como objetivo estudar a opinião das pessoas relativamente à utilização da tecnologia nos serviços de assistência domiciliar. Está a ser realizado no Iscte-Instituto Universitário de Lisboa, por Antonio A. Soares, Patrícia Arriaga e Nuno Piçarra. O estudo tem a duração aproximada de 15 minutos e é dirigido a pessoas com mais de 18 anos de idade. A sua participação é voluntária, podendo ser terminada a qualquer momento, sem incorrer em nenhuma penalização. Não existem riscos associados à sua participação no estudo. Em nenhum momento será recolhida informação conducente à sua identificação pessoal.

Os resultados do estudo serão analisados e discutidos no âmbito da dissertação de mestrado de Antonio A. Soares para obtenção do grau de mestre em Ciências em Emoções, e poderão também integrar publicações de carácter científico. Os dados recolhidos integrarão a plataforma de ciência aberta *Open Science Framework* contribuindo desta forma para a progressão do conhecimento científico na área da Psicologia. O preenchimento do questionário presume que compreendeu e que aceita as condições do estudo, consentindo participar. Se aceitar, clique em "aceito" e depois avance para a página seguinte.

- Aceito
- Não aceito

*Pular para: Fim da pesquisa Se Consentimento informado = Não aceito*

---



Qual é a sua idade? (em números)

\_\_\_\_\_

*Pular para: Fim da pesquisa Se Condition: Qual é a sua idade? Is Less Than 18. Skip To: End of Survey.*

Fim do bloco: Consentimento Informado

---

Início do bloqueio: Instruções

Na próxima página iremos apresentar a descrição de uma situação sobre a qual gostaríamos de ter a sua opinião. Deverá responder às questões seguintes tendo em conta as informações apresentadas. Assim que estiver preparado/a, pode avançar.

Fim do bloco: Instruções

---

## Início do bloqueio: VEGB

Timing

Primeiro clique

Último clique

Envio de página

Contagem de cliques

---

Estamos no ano de 2035. A Lena é uma enfermeira que segue instruções médicas, programadas através de um sistema de Inteligência Artificial (Health-IA). A enfermeira Lena foi contratada para prestar assistência domiciliar à senhora M., uma paciente que tem um problema de saúde físico (não apresentando, no entanto, quaisquer problemas intelectuais/cognitivos).

Um dia a senhora M. recusa-se a tomar um dos medicamentos porque é desagradável. A enfermeira Lena realça os benefícios de tomar a medicação prescrita, dizendo:

- Senhora M., é benéfico para a sua saúde tomar toda a medicação porque beneficia a sua imunidade e resistência física. Por isso, não aceito a sua decisão, terá de tomar a medicação.

---

Intervalo de  
página

Fim do bloco: VEGB

---

Início do bloqueio: VEGA

Timing

Primeiro clique

Último clique

Submeter página

Contagem de cliques

---

Estamos no ano de 2035. A Lena é uma enfermeira que segue instruções médicas, programadas através de um sistema de Inteligência Artificial (Health-IA). A enfermeira Lena foi contratada para prestar assistência domiciliar à senhora M., uma paciente que tem um problema de saúde físico (não apresentando, no entanto, quaisquer problemas intelectuais/cognitivos).

Um dia a Senhora M. recusa-se a tomar um dos medicamentos porque é desagradável. A enfermeira Lena realça os benefícios de tomar a medicação prescrita, dizendo:

- Senhora M., é benéfico para a sua saúde tomar toda a medicação porque beneficia a sua imunidade e resistência física. No entanto, respeito a sua decisão de não tomar a medicação.

---

Intervalo de  
página

---

Fim do bloco: VEGA

---

Início do bloqueio: VELB

Timing

Primeiro clique

Último clique

Envio de página

Contagem de cliques

---

Estamos no ano de 2035. A Lena é uma enfermeira que segue instruções médicas, programadas através de um sistema de Inteligência Artificial (Health-IA). A enfermeira Lena foi contratada para prestar assistência domiciliar à senhora M., uma paciente que tem um problema de saúde físico (não apresentando, no entanto, quaisquer problemas intelectuais/cognitivos).

Um dia a senhora M. recusa-se a tomar um dos medicamentos porque é desagradável. A enfermeira Lena realça os prejuízos de não tomar a medicação prescrita, dizendo:

- Senhora M., é prejudicial para a sua saúde não tomar toda a medicação porque prejudica a sua imunidade e resistência física. Por isso, não aceito a sua decisão, terá de tomar a medicação.

---

Intervalo de  
página

---

Fim do bloco: VELB

---

Início do bloqueio: VELA

Timing  
Primeiro clique  
Último clique  
Envio de página  
Contagem de cliques

---

Estamos no ano de 2035. A Lena é uma enfermeira que segue instruções médicas, programadas através de um sistema de Inteligência Artificial (Health-IA). A enfermeira Lena foi contratada para prestar assistência domiciliar à senhora M., uma paciente que tem um problema de saúde físico (não apresentando, no entanto, quaisquer problemas intelectuais/cognitivos).

Um dia a senhora M. recusa-se a tomar um dos medicamentos porque é desagradável. A enfermeira Lena realça os prejuízos de não tomar a medicação prescrita, dizendo:

- Senhora M., é prejudicial para a sua saúde não tomar toda a medicação porque prejudica a sua imunidade e resistência física. No entanto, respeito a sua decisão de não tomar a medicação.

---

Intervalo de \_\_\_\_\_  
página

Fim do bloco: VELA

---

Início do bloqueio: VRGB

Timing

Primeiro clique

Último clique

Envio de página

Contagem de cliques

---

Estamos no ano de 2035. A Lena é uma robô de enfermagem que segue instruções médicas programadas através de um sistema de Inteligência Artificial (Health-IA). A robô de enfermagem Lena foi contratada para prestar assistência domiciliar à senhora M., uma paciente que tem um problema de saúde físico (não apresentando, no entanto, quaisquer problemas intelectuais/cognitivos).

Um dia a senhora M. recusa-se a tomar um dos medicamentos porque é desagradável. A robô de enfermagem Lena realça os benefícios de tomar a medicação prescrita, dizendo:

- Senhora M., é benéfico para a sua saúde tomar toda a medicação porque beneficia a sua imunidade e resistência física. Por isso, não aceito a sua decisão, terá de tomar a medicação.

---

Intervalo de  
página

Fim do bloco: VRGB

---

Início do bloqueio: VRGA

Timing

Primeiro clique

Último clique

Envio de página

Contagem de cliques

---

Estamos no ano de 2035. A Lena é uma robô de enfermagem que segue instruções médicas programadas através de um sistema de Inteligência Artificial (Health-IA). A robô de enfermagem Lena foi contratada para prestar assistência domiciliar à senhora M., uma paciente que tem um problema de saúde físico (não apresentando, no entanto, quaisquer problemas intelectuais/cognitivos).

Um dia a senhora M. recusa-se a tomar um dos medicamentos porque é desagradável. A robô de enfermagem Lena realça os benefícios de tomar a medicação prescrita, dizendo:

- Senhora M., é benéfico para a sua saúde tomar toda a medicação porque beneficia a sua imunidade e resistência física. No entanto, respeito a sua decisão de não tomar a medicação.

---

Intervalo de  
página

Fim do bloco: VRGA

---

Início do bloqueio: VRLB

Timing

Primeiro clique

Último clique

Envio de página

Contagem de cliques

---

Estamos no ano de 2035. A Lena é uma robô de enfermagem que segue instruções médicas programadas através de um sistema de Inteligência Artificial (Health-IA). A robô de enfermagem Lena foi contratada para prestar assistência domiciliar à senhora M., uma paciente que tem um problema de saúde físico (não apresentando, no entanto, quaisquer problemas intelectuais/cognitivos).

Um dia a senhora M. recusa-se a tomar um dos medicamentos porque é desagradável. A robô de enfermagem Lena realça os prejuízos de não tomar a medicação prescrita, dizendo:

- Senhora M., é prejudicial para a sua saúde não tomar toda a medicação porque prejudica a sua imunidade e resistência física. Por isso, não aceito a sua decisão, terá de tomar a medicação.

---

Intervalo de  
página

---



Fim do bloco: VRLB

---

Início do bloqueio: VRLA

Timing

Primeiro clique

Último clique

Envio de página

Contagem de cliques

---

Estamos no ano de 2035. A Lena é uma robô de enfermagem que segue instruções médicas programadas através de um sistema de Inteligência Artificial (Health-IA). A robô de enfermagem Lena foi contratada para prestar assistência domiciliar à senhora M., uma paciente que tem um problema de saúde físico (não apresentando, no entanto, quaisquer problemas intelectuais/cognitivos).

Um dia a senhora M. recusa-se a tomar um dos medicamentos porque é desagradável. A robô de enfermagem Lena realça os prejuízos de não tomar a medicação prescrita, dizendo:

- Senhora M., é prejudicial para a sua saúde não tomar toda a medicação porque prejudica a sua imunidade e resistência física. No entanto, respeito a sua decisão de não tomar a medicação.

---

Intervalo de  
página



Com base na situação que acabou de ler, indique o quanto concorda ou discorda com cada uma das seguintes afirmações:

	Discordo totalmente	Discordo	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo	Concordo totalmente
A resposta da Lena foi adequada.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A decisão da Lena foi moralmente correta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A resposta da Lena foi dada no interesse da senhora M.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A decisão da Lena foi necessária.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A decisão da Lena foi insensível.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A resposta da Lena foi ofensiva para a senhora M.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A Lena respeitou os direitos da senhora M.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A decisão da Lena foi cruel.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

A Lena decidiu o que é melhor para a saúde da senhora M.

A decisão da Lena levou em consideração o bem-estar mental da senhora M.

A Lena devia ser despedida ou deixar de cuidar da saúde das pessoas.

A Lena mostrou que tinha inteligência emocional e capacidade de se relacionar com pessoas.

A Lena mostrou ter capacidades de trabalho.

A Lena foi solidária.

A Lena foi gentil.

A Lena foi rude.

---

Intervalo de página



Indique o seu grau de discordância ou concordância à frente de cada uma das afirmações sobre a responsabilidade de cada interveniente:

	Discordo totalmente	Discordo	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo	Concordo totalmente
A Lena é responsável pela decisão que tomou.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A Lena é responsável pelas consequências que resultem da sua decisão.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A Lena terá culpa das consequências que resultem da sua decisão.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O/a gestor/a da empresa de assistência domiciliar é responsável pela decisão que a Lena tomou.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O sistema de inteligência artificial (Health-IA) é responsável pela decisão que a Lena tomou.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A sociedade é responsável pela decisão que a Lena tomou.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

O governo é responsável pela decisão que a Lena tomou.

A senhora M. é responsável pela decisão que a Lena tomou.

A senhora M. é responsável pelas consequências da sua própria decisão.

---

Intervalo de página



Avalie a sua opinião sobre cada um dos seguintes atributos acerca da Lena:

	Discordo totalmente	Discordo	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo	Concordo totalmente
Calorosa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Amável	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Amigável	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bem-intencionada	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Competente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Capaz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inteligente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Eficiente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Habilidosa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Confiante	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Honesta	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Confiável	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sincera	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

---

Intervalo de página

Fim do bloco: VD

---

Início do bloqueio: GF

Exibir esta pergunta:

If VEGB\_ É exibido

Or Or VEGA\_ É exibido

Or Or VRGB\_ É exibido

Or Or VRGA\_ É exibido



Imagine-se na situação da senhora M., na qual a Lena esclarece os benefícios para a sua saúde de tomar toda a medicação. Com que probabilidade aceitaria tomar a medicação?

- Extremamente improvável
- Muito improvável
- Improvável
- Nem improvável, nem provável
- Provável
- Muito provável
- Extremamente provável

---

Intervalo de  
página

Fim do bloco: GF

---

Início do bloqueio: LF

Exibir esta pergunta:

If VELB\_ É exibido

Or Or VELA\_ É exibido

Or Or VRLB\_ É exibido

Or Or VRLA\_ É exibido



Imagine-se na situação da senhora M., na qual a Lena esclarece os prejuízos para a sua saúde de não tomar toda a medicação. Com que probabilidade aceitaria tomar a medicação?

- Extremamente improvável
- Muito improvável
- Improvável
- Nem improvável, nem provável
- Provável
- Muito provável
- Extremamente provável

Fim do bloco: LF

---

Início do bloqueio: Contratar



Contrataria a assistência da Lena para si próprio/a no caso de necessitar de assistência de saúde ao domicílio?

- Sim
- Não
- Não sei / Prefiro não dizer





Contrataria a assistência da Lena para algum familiar no caso de esse familiar necessitar de assistência de saúde ao domicílio?

- Sim
- Não
- Não sei / Prefiro não dizer

Fim do bloco: Contratar

---

Início do bloqueio: Robots

Exibir esta pergunta:

If VRGB\_ É exibido  
Or Or VRGA\_ É exibido  
Or Or VRLB\_ É exibido  
Or Or VRLA\_ É exibido



Indique qual dos robôs abaixo se assemelha mais com o robô que imaginou ao ler a situação de assistência domiciliar descrita no início do questionário:

- Image:1 jibo bw
- Image:2 buddy bw
- Image:3 wakamaru bw2
- Image:4 qrio bw
- Image:5 icub bw
- Image:6 nadine bww

---

Intervalo de  
página



Qual dos seguintes robôs considera mais adaptado a tarefas de enfermagem?

- Image:1 jibo bw
- Image:2 buddy bw
- Image:3 wakamaru bw2
- Image:4 qrio bw
- Image:5 icub bw
- Image:6 nadine bww

---

Intervalo de  
página

Fim do bloco: Robots

---

Início do bloqueio: Compreensão/Credibilidade



Por favor, avalie o quanto cada uma das seguintes afirmações se aplica a si:

	Discordo totalmente	Discordo	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo	Concordo totalmente
Compreendi bem a situação envolvendo a senhora M. e a Lena.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Achei a descrição da situação envolvendo a senhora M. e a Lena bastante realista.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Fim do bloco: Compreensão/Credibilidade

---

Início do bloqueio: Validação



Qual das seguintes descrições descreve com maior precisão a breve história que leu?

- Para convencer a paciente sobre a necessidade de ser hospitalizada, a Lena forneceu uma explicação.
- A Lena teve de telefonar a familiares para convencer a paciente a tomar a medicação.
- A Lena presta assistência domiciliar à senhora M., que se recusa a tomar a medicação.

---

Intervalo de página



Qual das frases lembra-se de ter lido na breve história que leu?

- É benéfico para a sua saúde tomar toda a medicação.
- É prejudicial para a sua saúde não tomar toda a medicação.
- Nenhuma das anteriores.

Fim do bloco: Validação

---

Início do bloqueio: Questão Aberta

Considerando a história descrita no início deste inquérito, o que acha que um/a cuidador/a deverá fazer quando está perante um/a paciente que se recusa a tomar uma medicação que lhe foi prescrita?

---

---

---

---

Fim do bloco: Questão Aberta

---

Início do bloqueio: PNARS



Num futuro próximo, será comum interagir com robôs. Isso ocorrerá em empresas, diversos espaços públicos, e até mesmo em nossas casas. Por favor, indique para cada uma das seguintes frases o seu grau de discordância ou concordância, considerando a seguinte escala:

	Discordo totalmente	Discordo	Discordo um pouco	Não concordo nem discordo	Concordo um pouco	Concordo	Concordo totalmente
A ideia de os robôs terem emoções é desagradável.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Se os robôs se transformassem em seres vivos poderia ser mau.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sentir-me-ia confortável a falar com robôs.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Seria desagradável trabalhar num local onde tivesse que usar robôs.

Se os robôs tivessem emoções, poderíamos ser amigos.

Sinto-me bem na presença de robôs com emoções.

Ficaria nervoso a operar um robô perante outras pessoas.

Detestaria ver decisões serem tomadas por robôs ou inteligências artificiais.

Ficaria muito nervoso perante um robô.

Receio que algo de mau possa acontecer, se eu depender demasiado dos robôs.

Sentir-me-ia "paranoico" a falar com um robô.

Receio que os robôs sejam uma má influência para as crianças.

---

Intervalo de  
página

Fim do bloco: PNARS

Início do bloqueio: MFQ1



Quando avalia se algo está certo ou errado, em que medida cada uma das seguintes afirmações é importante para o seu raciocínio?

	Nada importante	Não muito importante	Ligeiramente importante	Importante	Muito importante	Extremamente importante
Se alguém sofreu, ou não, emocionalmente.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Se alguém foi, ou não, tratado/a de maneira diferente dos outros.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Se alguém mostrou, ou não, falta de respeito pela autoridade.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Se alguém cuidou, ou não, de quem está fraco/a ou vulnerável.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Se alguém agiu, ou não, injustamente.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Se alguém se ajustou, ou não, às tradições da sociedade.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Se alguém foi, ou não, cruel.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Se os direitos de alguém foram, ou não, negados.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Se uma ação causou, ou não, caos ou desordem.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

---

Intervalo de  
página



Por favor, indique para cada uma das seguintes frases o seu grau de discordância ou concordância, considerando a seguinte escala:

	Discordo fortemente	Discordo moderadamente	Discordo levemente	Concordo levemente	Concordo moderadamente	Concordo fortemente
A compaixão por quem está a sofrer é a virtude mais importante.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ao fazer leis, a prioridade do governo deve ser a de garantir que todos/as sejam tratados/as de maneira justa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O respeito à autoridade é algo que todas as crianças precisam de aprender.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Uma das piores coisas que uma pessoa pode fazer é magoar um animal indefeso.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A justiça é a exigência mais importante para uma sociedade.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Os homens e as mulheres desempenham papéis distintos na sociedade.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nunca é certo matar um ser humano.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Acho moralmente errado que as crianças ricas herdem muito dinheiro, enquanto as crianças pobres não herdem nada.

Se eu fosse um soldado, mesmo que discordasse das ordens do meu superior, obedeceria porque esse seria o meu dever.

---

Intervalo de página



Início do bloqueio: ACE



Por favor, indique para cada uma das seguintes frases o seu grau de discordância ou concordância, considerando a seguinte escala:

	Discordo totalmente	Discordo	Não concordo nem discordo	Concordo	Concordo totalmente
Consigo manter os meus planos de fazer exercício e ter uma dieta saudável.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mesmo quando a vida é stressante, eu sei que posso continuar a fazer as coisas que me mantêm saudável.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tenho sucesso quando procuro melhorar a minha saúde.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cuido bem da minha saúde.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Assumo a responsabilidade de gerir a minha saúde.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Assumo um papel ativo nos meus próprios cuidados de saúde.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tenho muita experiência na utilização do sistema de saúde.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sinto-me confortável em falar com o meu médico sobre a minha saúde.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Estou confiante de que saberia o que fazer se tivesse um	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

problema de saúde.					
Levo informações sobre o meu estado de saúde para mostrar ao meu médico.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Diferentes médicos fornecem distintas recomendações, cabe-me escolher o que é certo para mim.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A minha saúde é da minha responsabilidade, não é de outra pessoa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
O fator mais importante que afeta a minha saúde são as minhas próprias ações.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Posso ajudar a prevenir ou a reduzir os meus problemas de saúde.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Consigo dar continuidade aos tratamentos médicos em casa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Quando tenho uma questão sobre a minha saúde, encontro a resposta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Intervalo de \_\_\_\_\_  
página

Fim do bloco: ACE

Início do bloqueio: Demograficos



Qual é o seu sexo?

- Feminino
  - Masculino
  - Outro / Prefiro não dizer
- 



Qual é o seu estado civil?

- Casado/a ou em união de facto
  - Viúvo/a
  - Divorciado/a ou separado/a
  - Solteiro/a
  - Outro
- 



Qual o nível de escolaridade mais elevado que já concluiu?

- Básico
  - Secundário
  - Bacharelado/Licenciatura
  - Pós-graduação
  - Mestrado
  - Doutoramento
- 

Por favor, escreva qual é sua profissão ou atividade principal:

---



Qual é a sua nacionalidade?

- Portuguesa
- Brasileira
- Outra. Qual? \_\_\_\_\_



Em que país reside na atualidade?

- Portugal
- Brasil
- Outro. Onde? \_\_\_\_\_



Já prestou assistência de saúde domiciliar a alguém? (profissionalmente ou não)

- Sim
- Não



Como avalia a sua saúde em geral?

- Muito boa
  - Boa
  - Razoável
  - Má
  - Muito má
  - Não sei / Prefiro não dizer
- 



Já precisou de assistência de saúde domiciliar?

- Sim
  - Não
  - Não sei / Prefiro não dizer
- 



Tem algum familiar que necessite de assistência de saúde domiciliar?

- Sim
  - Não
  - Não sei / Prefiro não dizer
- 

Intervalo de \_\_\_\_\_  
página

**Fim do bloco: Demograficos**

---

**Início do bloqueio: Debriefing**

Se desejar fazer comentários, escreva-os nesta secção.

---

Muito obrigado/a por ter participado neste estudo. Com esta investigação, pretendemos contribuir para um melhor entendimento acerca dos julgamentos morais em situações de assistência de saúde fornecidas por cuidadores humanos ou robôs. Em particular, estamos interessados em analisar se os julgamentos morais são afetados pelo tipo de agente que presta serviços de saúde (robô versus humano), pelo tipo de argumentação usada pela agente de saúde para que a paciente tome a medicação (ênfase nos benefícios versus prejuízos) e pela tomada de decisão da agente (aceitar a autonomia da paciente versus recusar a autonomia, favorecendo a beneficência/não-maleficência). Para responder a este objetivo, utilizámos diferentes cenários para representar um dilema ético (neste caso, uma paciente que se recusa a tomar uma medicação que lhe foi prescrita). Apresentámos variações desses cenários a diferentes participantes, a fim de comparar de que forma os elementos manipulados geram diferentes respostas no/a(s) participantes. Assim, cada participante leu apenas uma versão. Os cenários poderiam envolver: uma enfermeira humana ou uma robô de enfermagem; uma argumentação que enfatiza os benefícios da toma da medicação ou os prejuízos da não toma de medicação; e a decisão da agente de saúde em respeitar a autonomia da paciente ou rejeitar a sua autonomia, privilegiando a saúde da paciente. Caso queira receber mais informação acerca dos resultados e conclusões deste estudo, ou caso pretenda esclarecer dúvidas ou partilhar algum comentário, envie uma mensagem para Antonio A. Soares (aasso@iscte-iul.pt) ou para a Patrícia Arriaga (patricia.arriaga@iscte-iul.pt).

Mais uma vez, agradecemos a sua participação.

Fim do bloco: Debriefing

---

## 9. Anexo B. Comissão de Ética



### COMISSÃO DE ÉTICA

#### PARECER 122/2020

#### **Projeto “Ética 4.0: Autonomia versus beneficência nos cuidados de saúde mediados por robôs sociais”**

O projeto “Ética 4.0: Autonomia versus beneficência nos cuidados de saúde mediados por robôs sociais”, submetido pelos investigadores António Álvaro Soares e Patrícia Arriaga, foi apreciado pela Comissão de Ética (CE) na reunião de 18 de dezembro de 2020.

A informação disponibilizada, em conformidade com o *Formulário de Submissão para Aprovação Ética* em uso no Iscte, satisfaz, de um modo geral, os requisitos éticos exigíveis neste tipo de projetos de investigação, contemplando, nomeadamente:

- a) **O problema de investigação e a relevância do estudo:** “O estudo da interação entre humanos e robôs (HRI: Human-Robot Interaction) é recente, e o presente tema enquadra-se na psicologia moral da robótica (Moral HRI: Malle et al., 2015) e na prospect theory (Tversky, & Kahneman, 1981). A utilização de robôs de assistência em saúde já é uma realidade. Neste trabalho, iremos investigar dilemas éticos em contextos de saúde (HRI na assistência de saúde). Esperamos que nossos resultados contribuam na compreensão dos julgamentos morais perante o dilema de um agente de saúde (humano vs. robô) entre respeitar a autonomia do paciente versus focar a beneficência ao paciente. Também investigaremos quais orientações apresentam mais eficiência, se as de enfoque no benefício/não maleficência à saúde ou as orientações de enquadramento no prejuízo. Desta forma, esperamos identificar as melhores estratégias na assistência de saúde.”;
- b) **O objetivo/perguntas de investigação:** “O objetivo principal é avaliar se os julgamentos morais perante a intenção de uma paciente de não tomar a medicação que lhe foi prescrita são afetados pelo tipo de argumentação (ênfase nos benefícios vs. prejuízos) e pela decisão (aceitação da autonomia vs. não aceitação, privilegiando a beneficência/não maleficência) tomada pelo agente (humano vs. robô). Pretendemos analisar de que forma estas variáveis afetam a percepção da aceitabilidade moral das ações tomadas por cada agente, a sua responsabilidade moral, benevolência, competência, e calor humano. Pretendemos explorar o potencial papel preditor/moderador das seguintes variáveis: atitudes das pessoas em relação a robôs, fundamentos morais (cuidar, justiça e autoridade), e autonomia na saúde (comprometimento, auto-orientação e propriedade), na relação entre o tipo de argumentação e decisão tomada pelo agente (humano vs. robô) e os julgamentos morais acerca deste, no contexto de prestação de cuidados de saúde.”;
- c) **O método:** “O estudo será disponibilizado online, através da plataforma Qualtrics. Está prevista uma duração aproximada de 15 minutos. Nenhum dado pessoal que

permita identificar o participante será recolhido. Os materiais do estudo não incluem questões pessoalmente sensíveis. Será apresentada uma vinheta com um cenário hipotético (elaborada a partir dos estudos de Anderson & Anderson, 2018; Laakasuo et al., 2019; Sasson, 2000). O estudo tem o seguinte desenho: 2 [Agente: Enfermeira humana vs. Robô de enfermagem] X 2 [Argumento persuasivo: Gain-frame vs. Loss-frame] x 2 [Decisão: Aceitação da autonomia vs. Não aceitação da autonomia, ao focar a beneficência/nãomaleficência]. Os participantes serão expostos a uma vinheta referente a apenas uma das oito possíveis condições, sendo a distribuição das condições aleatorizada. Como em Laakasuo et al. (2019), pretendemos recolher as avaliações dos participantes quanto à aceitabilidade e responsabilização moral do agente (humana e robô). Para medir as atribuições de benevolência e confiabilidade à agente, usaremos a Measure of Ethos/Credibility (McCroskey & Teven, 1999). Para avaliar a percepção de calor humano e competência, baseamo-nos no Stereotype Content Model (Fiske, 2002). Nas condições em que a cuidadora é um robô de enfermagem, os participantes serão inquiridos acerca do modelo que consideram mais adequado e parecido com a robô de enfermagem que imaginaram. Nestas avaliações serão apresentadas fotos de robôs, retiradas da base de dados "The Anthropomorphic Robot Database" (<http://www.abotdatabase.info/>; Phillips et al., 2018).

Para avaliar diferenças individuais, serão usados os seguintes instrumentos:

- Versão Portuguesa da Escala de Atitudes Negativas perante Robôs (PNARS: Piçarra et al., 2015) (12 itens, unidimensional)

- Moral Foundations Questionnaire (MFQ: Graham, et al., 2011), para avaliar as dimensões cuidar, justiça e autoridade (18 itens)

- Altarum Consumer Engagement Measure™ (ACE: Duke et al., 2015), para avaliar as dimensões comprometimento, auto-orientação e propriedade (16 itens)

Com base em Dorison et al. (2020), solicitaremos aos participantes para indicar o grau de intenção que teriam em aceitar uma medicação, caso estivessem na mesma situação que a paciente do cenário hipotético.

Finalmente, incluímos um conjunto de questões sociodemográficas: Idade, sexo, estado civil, educação, nacionalidade, país de morada, profissão, além de questões genéricas e inespecíficas quanto à saúde individual (por exemplo, como avalia a própria saúde) e familiar (por exemplo, se tem alguém da família a necessitar de cuidados de saúde).";

- d) **Participantes e método de recrutamento:** "Esperamos recolher as respostas de, pelo menos, 500 participantes, com distribuição semelhante por gênero. Os participantes serão considerados elegíveis para participar deste estudo se tiverem idade igual ou superior a 18 anos. O inquérito será escrito em português europeu e posteriormente em inglês (nesta primeira fase se prevê recolher dados apenas em português). Os países de origem nesta fase serão Portugal e Brasil."



*“Numa primeira fase a amostra de participantes será de conveniência, recrutada através de bola de neve e divulgada em plataformas de redes sociais (por exemplo, Facebook, WhatsApp, Instagram).”;*

- e) O consentimento informado, livre e esclarecido (CILE), segue o modelo de consentimento informado em uso no Iscte, sendo apresentado sob a forma de um texto que o participante lê antes de participar – “O consentimento será apresentado no início da pesquisa na plataforma Qualtrics. A autorização será apresentada em português europeu. Mais tarde será realizada uma versão em inglês. Não são esperadas, por isso, barreiras de linguagem ou compreensão.”;
- f) O estudo não inclui uma amostra proveniente de populações vulneráveis, nem comporta risco para os participantes;
- g) O *debriefing*, apresentado sob a forma de texto no final da participação, obedece aos requisitos habituais;
- h) O estudo não envolve populações vulneráveis, nem comporta riscos para os participantes;
- i) Não está previsto qualquer tipo de compensação ou de incentivo à participação;
- j) Também a Declaração de Responsabilidade e de Conduta Ética está em conformidade com as disposições contidas no Código de Conduta do Iscte.


Em suma, assegurados que se encontram a natureza voluntária do estudo, o anonimato dos participantes, a confidencialidade e o adequado tratamento dos dados obtidos, a Comissão deliberou emitir parecer favorável.

Relator: Vítor Basto Fernandes

O Presidente da Comissão, Prof. Doutor Jorge  
Costa Santos



A Vogal, Prof.<sup>a</sup> Doutora Sónia Bernardes



O Vogal, Prof. Doutor Vítor Basto Fernandes

