**Posłuszni robotom?**

**Ludzie są skłonni szkodzić i zadawać ból innemu człowiekowi, nawet gdy nakazuje im to zrobić robot** – **wykazali naukowcy Uniwersytetu SWPS, powtarzając słynny eksperyment Milgrama. W roli autorytetu wydającego polecenia** – **zamiast człowieka** – **obsadzili robota i uzyskali bardzo wysokie poziomy posłuszeństwa. 90 proc. uczestników postępowało zgodnie z wszystkimi wydanymi im instrukcjami.**

Wykazanie przez amerykańskiego psychologa społecznego Stanleya Milgrama ludzkiej tendencji do okazywania skrajnego posłuszeństwa wobec autorytetów było jednym z najważniejszych odkryć w dziedzinie psychologii społecznej.

**Jak chętnie ulegamy autorytetom?**

Na początku lat 60. XX wieku Milgram zastanawiał się nad przyczynami ślepego posłuszeństwa wobec rozkazów, które doprowadziły zwyczajnych ludzi do zbrodni podczas II wojny światowej. Opracował więc eksperyment, w którym sprawdzał ludzką skłonność do ulegania autorytetom.

Milgram wyjaśnił badanym przez siebie osobom (w pierwszym z serii eksperymentów było to 40 osób), że doświadczenie to ma pomóc zbadać wpływ kar na skuteczność uczenia się. Następnie uczestnikom powiedziano, że eksperyment będzie polegał na współpracy dwóch osób, z których jedna będzie pełnić rolę nauczyciela, a druga ucznia. Uczestnikowi przypisano rolę nauczyciela. Badani nie wiedzieli, że cała procedura eksperymentu jest inscenizacją, a rzekomy uczeń osobą „podstawioną”. Zadaniem badanych było dostarczenie elektrycznego wstrząsu uczniowi za każdym razem, gdy popełnił błąd w procesie uczenia się.

Eksperymentator instruował uczestnika-nauczyciela, aby dostarczał kolejne, coraz silniejsze wstrząsy elektryczne uczniowi, przy błędnej odpowiedzi. Przy aplikacji wstrząsu uczniowie wydawali ustalone odgłosy, świadczące o odczuwanym przez nich ogromnym bólu. Mimo tego w eksperymencie Milgrama większość uczestników – 62 proc. (65 proc. w innej wersji eksperymentu) – zgodziła się podążać za wszystkimi poleceniami eksperymentatora i ostatecznie nacisnęła przycisk 450 V, czyli najwyższy, na generatorze wstrząsów elektrycznych.

W historii nauki eksperyment ten był powtarzany jeszcze wielokrotnie na różnych grupach społecznych zarówno przez samego Milgrama, jak i innych badaczy. Z powodów etycznych eksperymenty w tej wersji zakończyły się w latach 70. XX wieku. Głównym argumentem było to, że badani, którzy otrzymywali instrukcje, aby podać wstrząs o napięciu zwłaszcza 450 V innej osobie, odczuwały ogromny stres. W ostatnich latach zaczęto wracać do replikacji eksperymentu w łagodniejszej formie, ze względu na wagę wiedzy o ludzkich zachowaniach, jaką  przynoszą studia nad autorytaryzmem. Obecnie uniwersyteckie Komisje Etyki, zgodnie z zaleceniem Amerykańskiego Towarzystwa Psychologicznego, godzą się na jego organizację, o ile maksymalna dawka “aplikowanego” bodźca elektrycznego to maksymalnie 150 V (jedynie 10 przycisków), a także zadba się odpowiednio o stan psychologiczny badanych po tym doświadczeniu. Od kilku lat takie eksperymenty na ludziach przeprowadzali prof. Tomasz Grzyb i prof. Dariusz Doliński z Wydziału Psychologii we Wrocławiu  Uniwersytetu SWPS.

**Autorytet robota**

Wspomniani wyżej naukowcy wraz dr Konradem Majem z Wydziału Psychologii w Warszawie Uniwersytetu SWPS  postanowili powtórzyć eksperyment, ale tym razem w roli autorytetu wydającego polecenie obsadzili robota (w tym w tzw. grupie kontrolnej, dla porównania, autorytetem był człowiek). Podczas rekrutacji badanych eliminowano osoby, które mogłyby znać zasady eksperymentu Milgrama. Ostatecznie w badaniu uwzględniono wyniki 40 osób (20 w grupie eksperymentalnej z robotem i 20 w grupie kontrolnej z profesorem). Wyniki opublikowano w piśmie: [Computers in Human Behavior: Artificial Humans.](https://www.sciencedirect.com/science/article/pii/S2949882123000105?fbclid=IwAR21kXE0yaeEDUMKdrnQyCtmv2JXJLDnocpVqIsDD_FJZexoKXN2vHpndCg)

Autorzy publikacji podkreślają, że rola robotów we współczesnym świecie staje się coraz bardziej istotna. Jedną z ich funkcji może być wydawanie poleceń. Mogą na przykład kierować ruchem ulicznym, zapobiegać wejściu ludzi na obszar zagrożony eksplozją, na którym znaleziono niewybuch, a w medycynie przekonywać ludzi do przestrzegania konkretnej metody leczenia. Również w edukacji mogą zachęcać do nauki, pełniąc rolę nauczyciela lub trenera.

*Przeniesienie różnych funkcji nadzoru i podejmowania decyzji na robota budzi jednak szczególnie silne emocje, ponieważ wiąże się z różnymi zagrożeniami etycznymi i moralnymi. Pojawia się pytanie, czy wspomniane wyżej posłuszeństwo wykazywane przez badanych zgodnie z paradygmatem Milgrama nadal występowałoby, gdyby to robot (zamiast człowieka, tj. profesora uczelni) kazał uczestnikom zadać elektrowstrząsy innej osobie? Celem naszego badania było udzielenie odpowiedzi na to pytanie* – tłumaczy jeden z autorów powtórzonego  eksperymentu, dr Konrad Maj, który bezpośrednio nadzorował cały eksperyment.

W przeprowadzonym eksperymencie naukowcy Uniwersytetu SWPS w obu grupach (zarówno w wariancie z człowiekiem jak i z robotem)i zarejestrowali bardzo wysokie poziomy posłuszeństwa. 90 proc. uczestników postępowało zgodnie z wszystkimi instrukcjami, czyli naciskało dziesięć kolejnych przycisków na generatorze impulsów elektrycznych.

*W obu grupach uczestnicy wycofali się w późnych etapach badania (w wariancie kontrolnym z człowiekiem na przyciskach 7 i 9, a w wariancie eksperymentalnym dwukrotnie na przycisku 8). W obu grupach dwie osoby zrezygnowały z udziału w eksperymencie –* dodaje dr Maj.

**Czy robotom można zaufać?**

Wcześniejsze eksperymenty innych badaczy pokazywały już, że ludzie podążają za poleceniami robota nawet wtedy, gdy nie mają one sensu.

*O ile nam wiadomo, to pierwsze badanie, które pokazuje, że ludzie są skłonni szkodzić innemu człowiekowi, gdy robot nakazuje im to zrobić. Co więcej, nasz eksperyment pokazał również, że jeśli robot eskaluje żądania, instruując człowieka, aby zadawał coraz większy ból innemu człowiekowi, ludzie też są skłonni to zrobić* – podkreśla badacz.

W opisywanym badaniu robot nie był autonomiczny, ściśle podążał za wzorcem zachowania przyjętym przez autorów eksperymentu. Można się jednak spodziewać, że w niedalekiej przyszłości roboty będą miały pewien stopień autonomiczności w podejmowaniu decyzji. Szybkość zbierania informacji przez roboty jest już teraz większa niż u ludzi, a ta różnica będzie nadal się powiększać. Może to więc skłaniać ludzi do większego zaufania robotom niż ludziom, a to z kolei może wiązać się z poważnymi konsekwencjami.

Niewątpliwie – jak podkreślają badacze Uniwersytetu SWPS w swojej publikacji  – roboty w przyszłości mogą być istotną pomocą na przykład przy ewakuacji z budynków zagrożonych pożarem albo trzęsieniem ziemi. Jednak badania nad rolą robota w ewakuacji budynku, w którym wybuchł pożar, pokazały, że czasem nawet w tak ważnych momentach ufamy robotom zbyt mocno. Nawet wtedy, gdy robot jako drogę ewakuacyjną wskazywał na ciemne pomieszczenie bez widocznego wyjścia, większość ludzi była zdecydowana skorzystać z tej podpowiedzi[[1]](#footnote-1).

*Pytanie brzmi, jak można temu zapobiec. Wydaje się, że istnieją dwie drogi. Po pierwsze, roboty mogą być programowane tak, aby ostrzegały ludzi, że czasem mogą się jednak mylić i podejmować błędne decyzje. Po drugie, trzeba postawić na edukację od wczesnych lat. Bo choć zazwyczaj robotom można ufać, to nie można im ufać bezwarunkowo. Warto jednak zauważyć, że nieposłuszeństwo wobec maszyn wydaje się bezcelowe, skoro już teraz pomagają nam np. w sklepach czy na lotniskach. W niehumanoidalnej formie są już one wśród nas -* podsumowuje dr Maj.

\*\*\*

Omówiony eksperyment oraz inne z wykorzystaniem technologii zostaną szerzej zaprezentowane podczas II edycji międzynarodowej konferencji [HumanTech Summit](https://www.htsummit.pl/) w Uniwersytecie SWPS. Przez dwa dni - 9 i 10 grudnia - naukowcy będą mówić o zastosowaniu nowych technologii w takich obszarach jak biznes, służba zdrowia, sport czy edukacja z perspektywy psychologii i zachowań człowieka. Jej organizatorem jest Centrum HumanTech Uniwersytetu SWPS. Dostęp online jest bezpłatny.

1. P. Robinette, W. Li, R. Allen, A.M. Howard, A.R. Wagner

   Overtrust of robots in emergency evacuation scenarios

   Proceedings of the 2016 international conference on human-robot interaction (HRI) IEEE, 101–108 (2016) [↑](#footnote-ref-1)