

EFEKT PRZEDŁUŻONEGO WPŁYWU DEZINFORMACJI TEORIE, METODY BADANIA I WYZNACZNIKI WIELKOŚCI

Mikołaj Buczel¹, Adam Siwiak², Malwina Szpitalak³

CONTINUED INFLUENCE EFFECT OF MISINFORMATION.
THEORIES, RESEARCH METHODS AND DETERMINANTS

Summary. The continued influence effect (CIE) is a phenomenon in which certain information, although retracted and corrected, still affects one's reports, reasoning and decisions. A practical example of such phenomenon is the widespread myth that vaccines cause autism – therefore CIE research is of great practical importance as it provides insight into the psychological mechanisms of relying on and spreading misinformation. The aim of this article is to present research on the continued influence effect. The experimental methods of CIE research were discussed, as well as theories describing the cognitive mechanisms of undergoing misinformation. The most important findings in this field are also cited, including individual and procedural magnitude determinants of misinformation reliance and techniques to increase the effectiveness of corrections.

Key words: misinformation, continued influence effect, mental models, selective retrieval, misinformation reliance determinants

Wprowadzenie

Istnieją dobrze udokumentowane przypadki, w których pewne przekonanie utrzymuje się, mimo że zostało oficjalnie ogłoszone, że jest ono błędne. Chyba najbardziej charakterystycznym tego przykładem jest szeroko rozpowszechniony mit, że szczepionki powodują występowanie wśród dzieci autyzmu (Rao, Andrade, 2011). Przypadki tego typu nazwać można dezinformacją (Lewandowsky, Ecker,

¹ Instytut Psychologii, Uniwersytet Jagielloński w Krakowie (Institute of Psychology, Jagiellonian University in Krakow), ORCID: 0000-0002-9133-7930.

² Instytut Psychologii, Uniwersytet Jagielloński w Krakowie (Institute of Psychology, Jagiellonian University in Krakow).

³ Instytut Psychologii, Uniwersytet Jagielloński w Krakowie (Institute of Psychology, Jagiellonian University in Krakow), ORCID: 0000-0003-4928-5556.

Adres do korespondencji: Mikołaj Buczel,
e-mail: mikolaj.buczel@doctoral.uj.edu.pl

Cook, 2017). Choć w języku polskim termin ten ma najprawdopodobniej pochodzenie rosyjskie (*дезинформация; dezinformatsiya*) i w tym ujęciu oznacza celowe rozpowszechnianie fałszywych informacji z zamiarem wywarcia wpływu, w literaturze anglojęzycznej wykorzystuje się także szersze określenie *misinformation*, oznaczające po prostu nieprawdziwą lub niedokładną informację (patrz: Buczel, Siwiak, Szpitalak, 2020, dla rozróżnienia). Obydwa te pojęcia łączy jednak pewna specyficzna cecha – duża trudność w korygowaniu fałszywych treści. W walce z dezinformacją zalecane jest jak najszybsze jej wycofywanie, wyjaśniające powody dokonywanej korekty (Chan, Jones, Albarracín, 2017), jednak techniki przeciwdziałania dezinformacji rzadko okazują się skuteczne (Walter, Murphy, 2018). W niniejszym tekście przedstawiony zostanie efekt przedłużonego wpływu dezinformacji (*continued influence effect, CIE*), podstawowe informacje na temat tego zjawiska, metody jego badania, a także proponowane ujęcia teoretyczne tłumaczące poleganie na dezinformacji oraz czynniki zwiększające lub zmniejszające ku temu tendencje.

Efekt przedłużonego wpływu (CIE)

Efekt przedłużonego wpływu dezinformacji definiuje się jako ciągle wpływające na rozumowanie, wnioskowanie, pamięć i decyzje informacji, która, choć została wycofana, wciąż jest przez jednostkę powielana i utrwalana (Johnson, Seifert, 1994; Lewandowsky i in., 2012; Walter, Tukachinsky, 2020). Typowa procedura badania efektu przedłużonego wpływu obejmuje przedstawienie fikcyjnego scenariusza zdarzenia (np. pożaru hali magazynowej, kradzieży biżuterii lub wypadku autobusowego), a następnie wycofanie jednej z informacji zawartych w tej historii (Wilkes, Leatherbarrow, 1988). Choć wydawać by się mogło, że, po dostrzeżeniu dementującej ich dotychczasową wiedzę informacji, badani powinni formułować wnioski niezawierające dezinformacji, tak się jednak nie dzieje, a wycofanie (czyli zanegowanie lub korekta dezinformacji) jedynie w niewielkim stopniu redukuje poleganie na dezinformacji (np. Wilkes, Leatherbarrow, 1988; van Oostendorp, Bonbakker, 1999; Wilkes, Reynolds, 1999; Ecker, Lewandowsky, Tang, 2010; Ecker, Lewandowsky, Apai, 2011a; Ecker i in., 2011b; 2014a; Guillory, Geraci, 2013; Rich, Zaragoza, 2016; 2020; Ecker, Hogan, Lewandowsky, 2017; Ecker, Ang, 2019). Korekta bywa również całkowicie nieskuteczna (Johnson, Seifert, 1994; Ecker, Antonio, 2021), a w niektórych przypadkach może nawet spowodować paradoksalny wzrost polegania na dezinformacji (Nyhan, Reifler, 2010).

Niekiedy ludzie ulegają dezinformacji nawet przy niezwykle jasnych, bezpośrednich dowodach na jej fałszywość. W badaniu Jones i wsp. (2017) prezentowano uczestnikom opis sprzeczki między obywatelem a funkcjonariuszem policji, a następnie zawierający dezinformację raport policyjny na temat tego zdarzenia, w którym fałszywie obwiniono obywatela i usprawiedliwiono policjanta. Na koniec badani oglądali nagranie wideo z noszonej przez tego samego funkcjonariusza kamery, z którego jasno wynikało, że tak naprawdę to on zaatakował przechodnia.

Okazało się, że, pomimo tego ostatniego dowodu, badani zrzucali winę za całe zajście na obywatela. Co szczególnie niepokojące, skuteczne wydają się tylko te korekty, które pochodzą ze źródła zaufanego, lecz niekoniecznie fachowego (Guillory, Geraci, 2013; Pluviano, Della Sala, Watt, 2020; Ecker, Antonio, 2021). Może być to szczególnie niebezpieczne w sytuacji masowego rozprzestrzeniania dezinformacji przez różnego rodzaju osoby publiczne, takie jak aktorzy, celebryci lub sportowcy promujący w mediach społecznościowych szkodliwe teorie spiskowe, sprzeciwiając się konsensusowi naukowemu (np. łączące sieć 5G i rozprzestrzenianie COVID-19).

Teorie efektu przedłużonego wpływu

W literaturze proponuje się zazwyczaj dwie teorie poznawczych mechanizmów leżących u podłoża CIE: teorię modeli mentalnych (np. Seifert, 2002), w której zakłada się niepowodzenie w aktualizowaniu i integracji informacji w przyczynowym modelu poznawczym danego zdarzenia, oraz koncepcję selektywnego wydobywania informacji z pamięci poprzez poleganie na procesach automatycznych i przy jednoczesnym zawodzeniu procesów kontrolnych (np. Ecker, Swire, Lewandowsky, 2014b). Obydwie teorie zdaje się łączyć wybiórczość na poszczególnych etapach procesu pamięciowego (Hankała, 2001). Wybiórczość kodowania wydaje się kluczowym czynnikiem w teorii modeli mentalnych, gdzie wycofanie może nie zostać zakodowane w obrębie modelu w wyniku rozmaitych czynników, takich jak schematy poznawcze, wiedza lub ograniczona uwaga. Wybiórcze wydobycie dezinformacji z pamięci może zaś wynikać z obecności odpowiednich wskazówek, dostępnych schematów poznawczych lub użyteczności sytuacyjnej określonej informacji.

Aktualizacja modeli mentalnych

W teorii modeli mentalnych zakłada się, że ludzie konstruują poznawcze modele rozwijających się wydarzeń, czyli reprezentacje świata na podstawie dostępnych informacji i wiedzy jednostki, zbudowanych na zasadzie relacji przyczynowo-skutkowych (Trabasso, van den Broek, 1985). Modele mentalne nie są statycznymi strukturami i podlegają modyfikacjom tak, by odzwierciedlać aktualny stan rzeczy i tym samym umożliwiać wyciąganie wniosków w przypadku braku informacji przy jednoczesnym zachowaniu całościowej spójności (Albrecht, O'Brien, 1993). Takie aktualizacje mogą mieć charakter przyrostowy, kiedy do modelu dodaje się drobne elementy, lub globalny, kiedy dochodzi do rekonstrukcji i powstania nowego modelu. W przypadku CIE wydaje się, że nie dochodzi do aktualizacji globalnej, a zamiast tego do wyciągania wniosków potrzebnych do zachowania raczej lokalnej spójności (Seifert, 2002).

Początkowo zakładano, że do CIE dochodzi, gdy ludzie nie potrafią dokonywać edycji wspomnień i formułują szybkie wnioski na podstawie (dez)informacji, zanim nastąpi jej korekta (Wilkes, Leatherbarrow, 1988). Johnson i Seifert (1994) wykazały

jednak, że badani są w stanie wyciągać błędne wnioski na podstawie dezinformacji, nawet jeśli korekta jest podawana bezpośrednio po niej, kiedy jednostka nie powinna móc dokonać żadnych wniosków na podstawie (dez)informacji. Badaczki, zamiast poprzednich interpretacji, zaproponowały, że dezinformacja jest utrzymywana i wykorzystywana, ponieważ odgrywa ona rolę przyczynową dla opisywanego zdarzenia – gdy zaś unieważni się centralną dla modelu informację, powstanie luka, która sprawia, że model jest niespójny. Rozwiązaniem tego problemu miałyby być utrzymywanie w modelu zdyskredytowanej dezinformacji, ponieważ zapełnia ona powstałą lukę przyczynową – jednostka woli zatem zachować model spójny, lecz niedokładny i nieprawdziwy niż model niespójny, ale dokładniejszy i prawdziwy.

Jednym z dowodów na słuszność teorii modeli mentalnych dla wyjaśniania CIE jest skuteczność zastosowania alternatywnej dla dezinformacji treści, która zapełnia w jej miejsce lukę w modelu i powoduje redukcję CIE (np. Johnson, Seifert, 1994; Ecker, Lewandowsky, Tang, 2010; Nyhan, Reifler, 2015; Rich, Zaragoza, 2016; Ecker, Hogan, Lewandowsky, 2017) lub jego eliminację (Ecker, Lewandowsky, Apai, 2011a). Wydaje się również, że liczba powiązań przyczynowych związanych z dezinformacją jest bez znaczenia – nawet jedna niewyjaśniona przez inne okoliczności cecha opisywanego wydarzenia skutkuje wystąpieniem CIE (Wilkes, Reynolds, 1999). Także przedstawianie dezinformacji w sposób niejawni sprawia, że taka dezinformacja jest trudniejsza do korekty, co dzieje się najprawdopodobniej w wyniku wzmożonego wnioskowania przyczynowego (Rich, Zaragoza, 2016). To angażowanie się w wyjaśnienia przyczynowe zostało szerzej przebadane przez Hamby i wsp. (2020). W ich badaniu dezinformację wprowadzano w postaci treści retrospektywnie wyjaśniających opisywaną w scenariuszu historię. W zależności od warunku w ramach scenariusza wyjaśniono opisane wydarzenia lub ich nie wyjaśniono. Wyniki pokazały, że badani pozbawieni wyjaśnień częściej korzystali z wprowadzanych później dezinformacji, co wspiera hipotezę, że uzupełniali oni lukę przyczynową w wybrakowanym modelu mentalnym, nawet jeśli dezinformacja nie pochodziła bezpośrednio z samej historii. Dowodów dla teorii modeli mentalnych dostarczają również badania z użyciem neuroobrazowania. Gordon i in. (2017), za pomocą fMRI, wykazali, że w warunku wycofania dezinformacji aktywność rejonów odpowiedzialnych za integrowanie informacji była znacząco mniejsza niż w przypadku, gdy dezinformacja nie była wycofywana. Także osoby cechujące się niższą pojemnością pamięci roboczej, odpowiadającej m.in. za zdolności do integrowania sprzecznych informacji, w większym stopniu ulegają dezinformacji (Brydges, Gignac, Ecker, 2018).

Teoria selektywnego wydobywania

Alternatywna wobec modeli mentalnych jest teoria selektywnego wydobywania. W ujęciu tym zarówno dezinformacja, jak i treść wycofania są przechowywane w pamięci jednocześnie, a do CIE dochodzi, gdy aktywowana jest dezinformacja

przy niedostatecznym jej tłumieniu (np. Ecker, Swire, Lewandowsky, 2014b). U podstaw tej teorii leży rozróżnienie procesów automatycznych i kontrolowanych (Jacoby, 1991). Te pierwsze polegają na bezkontekstowym rozpoznawaniu informacji napotkanych wcześniej o charakterze ogólnym, podczas gdy te drugie pozwalają na wydobycie kontekstu, źródła i prawdziwości informacji. Zakłada się (Ayers, Reder, 1998), że informacje w pamięci konkurują między sobą o aktywację, a dostępne dla jednostki są tylko te, które wygryają konkurencję. Jednocześnie asocjacje między informacjami i ich kontekst (np. źródło) nie są dla jednostki dostępne i musi ona wykorzystać procesy strategiczne, aby uzyskać do nich dostęp. Do CIE dochodzi więc, gdy dezinformacja wygrywa konkurencję w aktywacji i jest wydobywana automatycznie np. w wyniku obecności odpowiedniej wskazówki, a oparte na zasobach poznawczych procesy kontrolowane są zakłócane i zawodzą w wydobyciu korekty (Ecker, Swire, Lewandowsky, 2014b). Takie automatyczne wydobywanie może się opierać na efekcie zaznajomienia, który polega na tym, że informacje znajome są częściej postrzegane jako prawdziwe (Dechêne i in., 2010) – gdy więc pojawia się jakakolwiek wskazówka mająca powiązanie z dezinformacją lub też ona sama po raz kolejny, zwiększa to jej szansę na aktywację. Można również spekulować, że dezinformacja jest wydobywana selektywnie i wykorzystywana wtedy, kiedy w danym momencie jest istotna w związku z określonymi wymaganiami sytuacji (patrz: sytuacyjna gotowość pamięci; Hankała, 2021).

Dowodami na słuszność teorii selektywnego wydobywania są wyniki badań Eckera i in. (2010) nad ostrzeżeniem o wystąpieniu dezinformacji, z których wynika, że może ono wspomagać procesy strategiczne i przyczyniać się do redukcji CIE. Swire i wsp. (2017b), wskazali zaś, że zwiększenie poziomu szczegółowości korekt (pobudzających procesy kontrolowane i ułatwiających konkurencję z dezinformacją) owocowało bardziej trwałymi zmianami przekonań. Także wielokrotne przetwarzanie dezinformacji skutkuje większym jej uleganiem (Ecker i in., 2011b), co w niektórych przypadkach może prowadzić niekiedy nawet do paradoksalnego wzrostu tendencji do jej ulegania, gdy jest ona powtarzana podczas korekty, skutkując tzw. efektem ognia wstecznego (*backfire effect*; Schwarz i in., 2007; patrz jednak: Ecker, Lewandowsky, Chadwick, 2020b). Z drugiej strony wyniki wielu badań wskazują jednak, że powtarzanie dezinformacji przy wycofywaniu poprawia jej korygowanie (Ecker, Hogan, Lewandowsky, 2017).

Innym modelem teoretycznym, często łączonym z omówionym powyżej, jest model przetwarzania negacji. Zgodnie z koncepcją Gilberta (1991) informacje podczas kodowania są zawsze traktowane jako prawdziwe i mogą dopiero zostać uznane za fałszywe później, poprzez dołączenie do nich negacyjnego znacznika (*negation tag*). Proces ten wymaga jednak wykorzystania zasobów poznawczych (jest to proces kontrolowany), wobec czego może czasem zawodzić, skutkując tym, że dezinformacja będzie traktowana jako prawdziwa. Nawet jeśli jednak dezinformacja zostanie prawidłowo oznaczona jako fałszywa, do CIE również może dochodzić – dołączanie do informacji znacznika negacji nie prowadzi bowiem do fu-

zji w pojedynczą reprezentację, a jedynie do połączenia informacji rdzeniowej ze znacznikiem. Tym samym istnieje możliwość, że znacznik zostanie od informacji rdzeniowej odłączony, skutkując traktowaniem jej jako prawdziwej (Mayo, Schul, Burnstein, 2004). Wiele badań nad CIE, których wyniki wspierają teorię selektywnego wydobywania, uwzględnia również mechanizmy kodowania negacji (np. Ecker, Lewandowsky, Tang, 2010; Ecker i in., 2011b; Lewandowsky i in., 2012; Swire, Ecker, Lewandowsky, 2017b).

Czynniki pozapamięciowe

Choć obydwie teorie wyjaśniające CIE mają charakter poznawczy, a zwłaszcza pamięciowy, zauważa się, że same czynniki pamięciowe nie są wystarczające w wyjaśnianiu mechanizmów CIE we wszystkich przypadkach i że należy brać pod uwagę także czynniki motywacyjne: postawy i światopoglądy. Proponuje się np., by mechanizmy motywowanego rozumowania stały za uleganiem dezinformacji w sytuacji, gdy wiąże się ona z poglądami politycznymi i uprzedzeniami (Nyhan, Reifler, 2010; Ecker i in., 2014a; Thorson, 2016; Ecker, Ang, 2019), choć można zauważyć, że traktuje się je raczej jako mechanizmy wspomagające procesy pamięciowe, szczególnie proces selektywnego wydobywania określonych informacji. Może być również tak, że jednostka nie akceptuje wycofania, ponieważ jest ono dla niej z jakiegoś powodu niewiarygodne (O'Rear, Radvansky, 2020; Ecker, Antonio, 2021) lub nie pochodzi z godnego zaufania źródła (Guillory, Geraci, 2013). W niektórych sytuacjach, kiedy źródło wycofania jest mniej wiarygodne niż źródło dezinformacji, poleganie na dezinformacji może być nawet do pewnego stopnia racjonalne (Connor Desai, Pilditch, Madsen, 2020). Wyniki ostatnich badań sugerują ponadto istnienie pewnych dowodów na to, że czynniki motywacyjne mogą odgrywać kluczową rolę w powstawaniu CIE równoległe do czynników poznawczych, nawet jeśli dezinformacja jest neutralna i nie ma powiązań ze światem rzeczywistym – wycofania dezinformacji mogą rodzić napięcie motywacyjne, które jest następnie redukowane przez zwiększenie wiary w dezinformację (Susmann, Wegener, 2021).

Metody badania efektu przedłużonego wpływu

W nurcie badań nad CIE wyróżnić można dwa wiodące schematy eksperymentalne. Pierwszy taki schemat można nazwać **procedurą narracyjną**. Obejmuje ona przedstawienie badanym materiału w formie scenariusza dotyczącego jakiejś fikcyjnej historii, w zamyśle stylizowanego na prasowe wiadomości (np. Wilkes, Leatherbarrow, 1988; Johnson, Seifert, 1994; Ecker, Lewandowsky, Tang, 2010; Guillory, Geraci, 2013; Rich, Zaragoza, 2016). Przedstawiona historia może dotyczyć różnorodnych treści, m.in. pożaru hali magazynowej artykułów papierniczych, kradzieży biżuterii lub starającego się o reelekcję polityka, z reguły są to jednak historie fikcyjne, nieodnoszące się do wydarzeń ze świata rzeczywistego. Scenariusz w zdecydo-

wanej większości przypadków przedstawiany jest sekwencyjnie w formie od kilku do kilkunastu jedno- lub dwuzdaniowych fragmentów, do których osoba badana nie ma możliwości powrotu. W jednej z początkowych wiadomości prezentuje się dezinformację, czyli treść, która jest pod koniec scenariusza wycofywana – np. jeśli osoba badana najpierw zapoznaje się z informacją o znalezieniu na miejscu opisywanego pożaru substancji łatwopalnych, później dowiaduje się, że informacja ta nie była prawdziwa. Po zapoznaniu się z całością materiału badany prezentuje się test, zazwyczaj w formie pytań otwartych, wymagających zarówno podania zapamiętanych ze scenariusza faktów, jak i wnioskowania na ich podstawie (np. Johnson, Seifert, 1994; Ecker, Lewandowsky, Tang, 2010; Guillory, Geraci, 2013), choć można również napotkać inne formy testów, np. w postaci pytań zamkniętych lub skali szacunkowych (Nyhan, Reifler, 2015; Thorson, 2016; Connor Desai, Reimers, 2019). Od opisanego powyżej typowego schematu odchodzi się w części badań, prezentując dezinformację i wycofanie w odwróconej kolejności (Gordon, Ecker, Lewandowsky, 2019) lub prezentując treści bezpośrednio dotyczące realnych wydarzeń, gdzie całość materiału, włączając w to zarówno dezinformację, jak i wycofanie, była już wcześniej obecna w przestrzeni publicznej (Lewandowsky i in., 2005; Nyhan, Reifler, 2010). Istnieją także badania łączące narracyjną procedurę CIE z efektem dezinformacji po zdarzeniu (*post-event misinformation effect*, PME; Loftus, Miller, Burns, 1978), gdzie najpierw prezentuje się materiał oryginalny np. w formie filmu, po czym jego opis zawierający dezinformację, którą się następnie wycofuje (np. Crozier, Strange, 2019).

Drugi schemat procedur badawczych CIE można określić mianem **procedury dyskonfirmacyjnej** lub też prościej **obalaniem mitów**. W przeciwieństwie do procedury narracyjnej nie używa się scenariuszy, a zestawów zazwyczaj kilkudziesięciu krótkich twierdzeń (np. Swire i in., 2017a; Swire, Ecker, Lewandowsky, 2017b; Aird i in., 2018; Ecker, Butler, Hamby, 2020a; Ecker, Lewandowsky, Chadwick, 2020b; Ecker i in., 2020c). Wśród nich znajdują się zarówno twierdzenia prawdziwe, zgodne z aktualnym stanem wiedzy (fakty), jak i twierdzenia fałszywe (mity). Twierdzenia te dotyczą najczęściej wiedzy ogólnej, specjalistycznej i naukowej o niejasnej dla przeciętnego badanego prawdziwości (np. „Kłamcy czasami zdradzają się fizycznymi wskazówkami, takimi jak patrzeć w prawo lub niepatrzeć w oczy”) lub wiedzy politycznej, o bieżących wydarzeniach. Po przedstawieniu badanym mitów następuje ich wycofanie, często połączone z dodatkowym wyjaśnieniem, a fakty są potwierdzane (tzn. badanych po prostu informuje się, że jakaś treść wcześniej im prezentowana była prawdziwa). Test przeprowadza się dwukrotnie: (1) po prezentacji dezinformacji i przed wycofaniem oraz (2) po wycofaniu. Ma on w prawie każdym przypadku formę skali szacunkowej, gdzie badany ocenia swoją pewność wobec tego, na ile, według niego, dane stwierdzenie jest prawdziwe. Miara pretestu stanowi bazę do późniejszej oceny różnicy między jego wynikiem a wynikiem posttestu – mierzona jest więc zmiana wiary w dezinformację lub jej stabilność mimo korekty. Procedura ta w większym stopniu bada bezpośrednią wiarę w dezinformację niż wnioskowanie na jej podstawie.

Wyznaczniki wielkości CIE

Choć efekt przedłużonego wpływu jest łatwo replikowalny (patrz: Lewandowski i in., 2012; Walter, Tukachinsky, 2020), może przyjmować różne wielkości w zależności od rozmaitych czynników. Idąc za Polczykiem (2007), który to samo uczynił dla efektu dezinformacji po zdarzeniu (PME), proponujemy podział na wyznaczniki indywidualne i proceduralne.

Wyznaczniki indywidualne

Jak wspomniano wcześniej, wpływ dezinformacji może być większy w zależności od czynników utrudniających wykorzystywanie procesów kontrolowanych. Jednym z czynników ograniczających ich funkcjonowanie jest wiek – osoby starsze słabiej pamiętają fakty, lecz, jak się okazuje, są tak samo jak młodszy dorośli podatne na dezinformację w procedurze narracyjnej (Guillory, Geraci, 2010). Różnice między osobami w wieku średnim a podeszłym dało się natomiast zaobserwować dla obalania mitów – osoby starsze są mniej skłonne do korygowania swoich pierwotnych przekonań (Swire, Ecker, Lewandowski, 2017b). Także mała pojemność pamięci roboczej istotnie zwiększa szansę na uleganie dezinformacji (Brydges, Gignac, Ecker, 2018). Czynnikiem ograniczającym wpływ dezinformacji wydaje się natomiast sceptycyzm, zarówno sytuacyjny, jak i ten będący cechą osobowościową (Lewandowski i in., 2005; 2009). Sceptycyzm zdaje się skłaniać jednostkę do kwestionowania pochodzenia informacji, zanim zostanie ona uznana za nieprawdziwą i wycofana, przydzielając więcej zasobów poznawczych do oceny prawdziwości napływających treści.

Istotnym czynnikiem wpływającym na CIE jest motywowanie rozumowanie. Teoria motywowanego rozumowania zakłada, że sposób, w jaki jednostka przetwarza napływające doń informacje, zależy od jej motywacji – może być ona zmotywowana do wyciągania wniosków pasujących do jej wcześniejszych postaw i przekonań (Kunda, 1990). U osób o odmiennych przekonaniach te same informacje, nawet jeśli przedstawiane są neutralnie, powodują aktualizację przekonań w przeciwnych kierunkach (Jerit, Barabas, 2012). Motywowanie rozumowanie wydaje się w przypadku CIE szczególnie silne, gdy dezinformacja dotyczy treści światopoglądowych. Ecker i in. (2011a) prezentowali badanym scenariusz, w którym za katastrofę samolotu winą obarczono islamskich terrorystów, choć potem okazało się, że przyczyną była awaria. Wyniki pokazały, że osoby z wysokim poziomem uprzedzeń częściej nawiązywały do wątków terrorystycznych. Podobne wnioski można wyciągnąć z badań na próbie australijskiej, które dotyczyły stereotypowego postrzegania osób pochodzenia aborygeńskiego (Ecker i in., 2014a). Jeszcze mocniej podatność na dezinformację widać w przypadku poglądów politycznych. W tego typu badaniach mierzona jest zazwyczaj wiara w dezinformację u osób o skrajnie różnych poglądach (liberalnych i konserwatywnych). W zdecydowanej większości

przypadków wykazuje się, że postawy polityczne mają duży wpływ na przetwarzanie dezinformacji i wycofań (np. Thorson, 2016; Swire i in., 2017a; Aird i in., 2018; Ecker, Ang, 2019; Ecker, Sze, Andreotta, 2021), przy czym efekt ten jest zdecydowanie silniejszy wśród konserwatystów, u których korekta dezinformacji może nawet zwiększyć wiarę w nieprawdziwą informację w porównaniu do warunku braku korekty (Nyhan, Reifler, 2010). Interpretując wpływ motywowanego rozumowania na procesy odpowiedzialne za CIE uważa się, że wycofania niezgodne z nastawieniem jednostki są skuteczne w takim stopniu, w jakim nie generują one potrzeby zmiany postawy (Ecker i in., 2014a).

Wyznaczniki proceduralne

Wielkość CIE może zależeć także od warunków eksperymentalnych wprowadzanych przez badaczy. W procedurach badań wielkości CIE bardzo często zwracano uwagę na wyrazisty sposób prezentowania zarówno dezinformacji, jak i jej wycofania. Jeśli dezinformacja umieszczona była w centralnej – z punktu widzenia treści – części scenariusza, co czyniło ją bardziej wyrazistą, to badani częściej jej ulegali (Wilkes, Leatherbarrow, 1988). Uleganie dezinformacji oraz zaakceptowanie wycofania zależą też od siły ich zakodowania. Zaobserwowano, że wielokrotne wycofania nie redukowały wpływu dezinformacji w większym stopniu niż pojedyncze odwołanie, jednak kilkukrotne powtórzenie dezinformacji wymagało wielokrotnego jej wycofywania, aby znacząco zredukować CIE (Ecker i in., 2011b). Wyrazistość dezinformacji wiąże się także ze wspomnianym wcześniej efektem ognia wstecznego. Choć może on występować w przypadku dezinformacji politycznej, gdyż wiara w fałszywe doniesienia jest mocno zależna od poglądów ideologicznych (tzw. *worldview backfire effect*, czyli światopoglądowy efekt ognia wstecznego; Nyhan, Reifler, 2010), w przypadku neutralnego materiału CIE istnieje także inny jego rodzaj. Istnieją doniesienia, że możliwe jest występowanie efektu ognia wstecznego polegającego na zaznajomieniu (*familiarity backfire effect*), czyli zwiększonej znajomości dezinformacji w związku z jej powtarzalnością, co powoduje łatwiejsze jej wyszukiwanie w systemie pamięci (np. Schwarz i in., 2007). Ecker i in. (2020b) zauważają jednak, że dowody empiryczne na efekt ognia wstecznego polegającego na zaznajomieniu są mocno wątpliwe i ograniczają się do niewielkiej liczby badań.

Innym niezwykle ważnym czynnikiem jest wiarygodność. W przypadku CIE analizowano zwykle wiarygodność źródeł wycofania, biorąc pod uwagę dwa jego aspekty: fachowość i zaufanie (Pornpitakpan, 2004). Sama fachowość źródła (czyli kompetencje do wydawania trafnych osądów) nie jest w stanie zredukować polegania na dezinformacji, podczas gdy źródło zaufane (czyli postrzegane jako chętne do podania właściwych informacji) redukuje CIE niezależnie od stopnia jego fachowości (Guillory, Geraci, 2013; Pluviano, Della Sala, Watt, 2020; Ecker, Antonio, 2021). Może być również tak, że jednostka podtrzymuje wiarę w dezinformację kosztem

wiarygodności źródła wycofania niezależnie od tego, czy było ono takie samo, czy różne od pierwszego (Connor Desai, Pilditch, Madsen, 2020).

Na wielkość CIE wpływają także odstępy czasu między poszczególnymi etapami procedury badawczej. Swire i in. (2017b) wykazali, że po upływie tygodnia znacząco rosła wiara w poprawioną dezinformację w porównaniu do testowania natychmiastowego lub odroczonego o 30 minut. Było to jeszcze bardziej widoczne po trzech tygodniach, szczególnie w grupie osób starszych. Dowodów na to, że wpływ czasu ma wpływ na CIE w procedurze narracyjnej dostarczyły zaś wyniki badań Richa i Zaragozy (2020). Choć okazało się, że czas między wprowadzeniem dezinformacji a dostarczeniem wycofania nie ma większego wpływu na uleganie dezinformacji, to wzrost polegania na fałszywych treściach następował po upływie kilku dni od podania wycofania pomimo pamięci wystąpienia korekty. Badacze stwierdzili, że zmienność wpływu dezinformacji w czasie może wynikać z pewnej nieskuteczności aktualizowania modelu mentalnego – jeśli początkowy błędny model nie został zaktualizowany przez dodanie do niego korekty, mogło nie dojść do trwałej w czasie zmiany w strukturze wiedzy.

Ważne z perspektywy rozpowszechniania dezinformacji wydają się także czynniki na nią uodparniające. Jednym z nich jest, omówione wcześniej, podanie podczas wycofania alternatywnej dla dezinformacji treści, która zastępowałaby ją w obrębie sieci przyczynowej modelu mentalnego i ograniczała jej wpływ (np. Johnson, Seifert, 1994). Porównywalnie skuteczne wydaje się także zastosowanie ostrzeżenia przed możliwością wystąpienia dezinformacji, choć musi być ono odpowiednio szczegółowe; połączenie ostrzeżenia i alternatywy skutkuje zaś najsilniejszą redukcją CIE (Ecker, Lewandowsky, Tang, 2010). Obie te techniki są jednak z kilku względów problematyczne. Po pierwsze, ani alternatywa, ani ostrzeżenie, ani nawet ich połączenie nie są w stanie całkowicie wyeliminować wpływu dezinformacji (np. Johnson, Seifert, 1994; Ecker, Lewandowsky, Tang, 2010; Rich, Zaragoza, 2020). Bywa również, że alternatywa nie jest w żadnym stopniu skuteczna (van Oostendorp, Bonebakker, 1999; O'Rear, Radvansky, 2020). Ponadto obie te techniki nie wydają się mieć dużego zastosowania praktycznego, gdyż znalezienie wiarygodnej alternatywy w warunkach pozalaboratoryjnych często bywa niemożliwe (Lewandowsky i in., 2012), a ostrzeganie ludzi, zanim spotkają się z konkretną dezinformacją, jest raczej mało realne (Swire, Ecker, 2018). Wydaje się wobec tego, że konieczna jest dalsza eksploracja czynników uodparniających na efekt przedłużonego wpływu i to szczególnie takich o zastosowaniu praktycznym. W dobie tak szeroko rozpowszechnianej dezinformacji należy przedsięwziąć ku temu celowi szczególne wysiłki.

Literatura cytowana

Aird, M.J., Ecker, U.K.H., Swire, B., Berinsky, A.J., Lewandowsky, S. (2018). Does truth matter to voters? The effects of correcting political misinformation in an Australian sample. *Royal Society Open Science*, 5(12), 180593.

- Albrecht, J.E., O'Brien, E.J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(5), 1061–1070.
- Ayers, M.S., Reder, L.M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin & Review*, 5(1), 1–21.
- Brydges, C.R., Gignac, G.E., Ecker, U.K.H. (2018). Working memory capacity, short-term memory capacity, and the continued influence effect: A latent-variable analysis. *Intelligence*, 69, 117–122.
- Buczel, M., Siwiak, A., Szpitalak, M. (2020). Continued influence of misinformation effect in the context of forensic psychology. *Problems of Forensic Sciences*, 121, 5–19.
- Chan, M.P.S., Jones, C.R., Albarracín, D. (2017). Countering false beliefs: An analysis of the evidence and recommendations of best practices for the retraction and correction of scientific misinformation. W. K. Hall Jamieson, D. Kahan, D.A. Scheufele (red.), *The Oxford handbook of the science of science communication* (s. 341–349). Oxford University Press.
- Connor Desai, S.A., Pilditch, T.D., Madsen, J.K. (2020). The rational continued influence of misinformation. *Cognition*, 205, 104453.
- Connor Desai, S.A., Reimers, S. (2019). Comparing the use of open and closed questions for Web-based measures of the continued-influence effect. *Behavior Research Methods*, 51(3), 1426–1440.
- Crozier, W.E., Strange, D. (2019). Correcting the misinformation effect. *Applied Cognitive Psychology*, 33(4), 585–595.
- Dechêne, A., Stahl, C., Hansen, J., Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, 14(2), 238–257.
- Ecker, U.K.H., Ang, L.C. (2019). Political attitudes and the processing of misinformation corrections. *Political Psychology*, 40(2), 241–260.
- Ecker, U.K.H., Antonio, L.M. (2021). Can you believe it? An investigation into the impact of retraction source credibility on the continued influence effect. *Memory & Cognition*, 49, 631–644.
- Ecker, U.K.H., Butler, L.H., Hamby, A. (2020a). You don't have to tell a story! A registered report testing the effectiveness of narrative versus non-narrative misinformation corrections. *Cognitive Research: Principles and Implications*, 5(1), 1–26.
- Ecker, U.K.H., Hogan, J.L., Lewandowsky, S. (2017). Reminders and repetition of misinformation: Helping or hindering its retraction? *Journal of Applied Research in Memory and Cognition*, 6(2), 185–192.
- Ecker, U.K.H., Lewandowsky, S., Apai, J. (2011a). Terrorists brought down the plane!—No, actually it was a technical fault: Processing corrections of emotive information. *Quarterly Journal of Experimental Psychology*, 64(2), 283–310.
- Ecker, U.K.H., Lewandowsky, S., Chadwick, M. (2020b). Can corrections spread misinformation to new audiences? Testing for the elusive familiarity backfire effect. *Cognitive Research: Principles and Implications*, 5(1), 1–25.

- Ecker, U.K.H., Lewandowsky, S., Fenton, O., Martin, K. (2014a). Do people keep believing because they want to? Preexisting attitudes and the continued influence of misinformation. *Memory & Cognition*, 42(2), 292–304.
- Ecker, U.K.H., Lewandowsky, S., Swire, B., Chang, D. (2011b). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review*, 18(3), 570–578.
- Ecker, U.K.H., Lewandowsky, S., Tang, D.T. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition*, 38(8), 1087–1100.
- Ecker, U.K.H., O'Reilly, Z., Reid, J.S., Chang, E.P. (2020c). The effectiveness of short-format refutational fact-checks. *British Journal of Psychology*, 111(1), 36–54.
- Ecker, U.K.H., Swire, B., Lewandowsky, S. (2014b). Correcting misinformation – A challenge for education and cognitive science. W: D.N. Rapp, J.L.G. Braasch (red.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (s. 13–37). MIT Press.
- Ecker, U.K.H., Sze, B.K.N., Andreotta, M. (2021). Corrections of political misinformation: No evidence for an effect of partisan worldview in a US convenience sample. *Philosophical Transactions of the Royal Society B*, 376(1822), 20200145.
- Gilbert, D.T. (1991). How mental systems believe. *American Psychologist*, 46, 107–119.
- Gordon, A., Brooks, J.C., Quadflieg, S., Ecker, U.K.H., Lewandowsky, S. (2017). Exploring the neural substrates of misinformation processing. *Neuropsychologia*, 106, 216–224.
- Gordon, A., Ecker, U.K.H., Lewandowsky, S. (2019). Polarity and attitude effects in the continued-influence paradigm. *Journal of Memory and Language*, 108, 104028.
- Guillory, J.J., Geraci, L. (2010). The persistence of inferences in memory for younger and older adults: Remembering facts and believing inferences. *Psychonomic Bulletin & Review*, 17(1), 73–81.
- Guillory, J.J., Geraci, L. (2013). Correcting erroneous inferences in memory: The role of source credibility. *Journal of Applied Research in Memory and Cognition*, 2(4), 201–209.
- Hamby, A., Ecker, U.K.H., Brinberg, D. (2020). How Stories in Memory Perpetuate the Continued Influence of False. *Information Journal of Consumer Psychology*, 30(2), 240–259.
- Hankała, A. (2001). *Wybiórczość ludzkiej pamięci*. Wydawnictwa Uniwersytetu Warszawskiego.
- Hankała, A. (2021). *Sytuacyjna gotowość pamięci*. CeDeWu.
- Jacoby, L.L. (1991). A Process Dissociation Framework: Separating Automatic from Intentional Uses of Memory. *Journal of Memory and Language*, 30, 513–541.
- Jerit, J., Barabas, J. (2012). Partisan perceptual bias and the information environment. *Journal of Politics*, 74, 672–1684.
- Johnson, H.M., Seifert, C.M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6), 1420–1436.

- Jones, K.A., Crozier, W.E., Strange, D. (2017). Believing is seeing: Biased viewing of body-worn camera footage. *Journal of Applied Research in Memory and Cognition*, 6(4), 460–474.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498.
- Lewandowsky, S., Ecker, U.K.H., Cook, J. (2017). Beyond misinformation: Understanding and coping with the “post-truth” era. *Journal of Applied Research in Memory and Cognition*, 6(4), 353–369.
- Lewandowsky, S., Ecker, U.K.H., Seifert, C.M., Schwarz, N., Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3), 106–131.
- Lewandowsky, S., Stritzke, W.G.K., Oberauer, K., Morales, M. (2005). Memory for fact, fiction, and misinformation: The Iraq War 2003. *Psychological Science*, 16(3), 190–195.
- Lewandowsky, S., Stritzke, W.G.K., Oberauer, K., Morales, M. (2009). Misinformation and the ‘War on Terror’: When memory turns fiction into fact. W: W.G.K. Stritzke, S. Lewandowsky, D. Denmark, J. Clare, F. Morgan (red.), *Terrorism and torture: An interdisciplinary perspective* (s. 179–203). Cambridge University Press.
- Loftus, E.F., Miller, D.G., Burns, H.J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, 4(1), 19–31.
- Mayo, R., Schul, Y., Burnstein, E. (2004). “I am not guilty” vs “I am innocent”: Successful negation may depend on the schema used for its encoding. *Journal of Experimental Social Psychology*, 40(4), 433–449.
- Nyhan, B., Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2), 303–330.
- Nyhan, B., Reifler, J. (2015). Displacing misinformation about events: An experimental test of causal correction. *Journal of Experimental Political Science*, 2(1), 81–93.
- O’Rear, A.E., Radvansky, G.A. (2020). Failure to accept retractions: A contribution to the continued influence effect. *Memory & Cognition*, 48(1), 127–144.
- Pluviano, S., Della Sala, S., Watt, C. (2020). The effects of source expertise and trustworthiness on recollection: The case of vaccine misinformation. *Cognitive Processing*, 21, 321–330.
- Polczyk, R. (2007). *Mechanizmy efektu dezinformacji w kontekście zeznań świadka naocznego*. Wydawnictwo Uniwersytetu Jagiellońskiego.
- Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades’ evidence. *Journal of Applied Social Psychology*, 34(2), 243–281.
- Rao, T.S., Andrade, C. (2011). The MMR vaccine and autism: Sensation, refutation, retraction, and fraud. *Indian Journal of Psychiatry*, 53(2), 95–96.
- Rich, P.R., Zaragoza, M.A. (2016). The continued influence of implied and explicitly stated misinformation in news reports. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42, 62–74.

- Rich, P.R., Zaragoza, M.S. (2020). Correcting Misinformation in News Stories: An Investigation of Correction Timing and Correction Durability. *Journal of Applied Research in Memory and Cognition*, 9(3), 310–322.
- Schwarz, N., Sanna, L.J., Skurnik, I., Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology*, 39, 127–161.
- Seifert, C.M. (2002). The continued influence of misinformation in memory: What makes a correction effective? W: B.H. Ross (red.), *The psychology of learning and motivation: Advances in research and theory* (s. 265–292). Academic Press.
- Susmann, M.W., Wegener, D.T. (2021). The role of discomfort in the continued influence effect of misinformation. *Memory & Cognition*, 1–14.
- Swire, B., Berinsky, A.J., Lewandowsky, S., Ecker, U.K.H. (2017a). Processing political misinformation: Comprehending the Trump phenomenon. *Royal Society Open Science*, 4(3), 160802.
- Swire, B., Ecker, U.K.H. (2018). Misinformation and its correction: Cognitive mechanisms and recommendations for mass communication. W: B.G. Southwell, E.A. Thorson, L. Sheble (red.), *Misinformation and mass audiences* (s. 195–221). University of Texas Press.
- Swire, B., Ecker, U.K.H., Lewandowsky, S. (2017b). The role of familiarity in correcting inaccurate information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(12), 1948–1961.
- Thorson, E. (2016). Belief echoes: The persistent effects of corrected misinformation. *Political Communication*, 33(3), 460–480.
- Trabasso, T., van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, 24(5), 612–630.
- van Oostendorp, H., Bonebakker, C. (1999). Difficulties in updating mental representations during reading news reports. W: H. van Oostendorp, S.R. Goldman (red.), *The construction of mental representations during reading* (s. 319–330). Lawrence Erlbaum Associates.
- Walter, N., Murphy, S.T. (2018). How to unring the bell: A meta-analytic approach to correction of misinformation. *Communication Monographs*, 85(3), 423–441.
- Walter, N., Tukachinsky, R. (2020). A meta-analytic examination of the continued influence of misinformation in the face of correction: How powerful is it, why does it happen, and how to stop it? *Communication Research*, 47(2), 155–177.
- Wilkes, A.L., Leatherbarrow, M. (1988). Editing episodic memory following the identification of error. *The Quarterly Journal of Experimental Psychology*, 40(2), 361–387.
- Wilkes, A.L., Reynolds, D.J. (1999). On certain limitations accompanying readers' interpretations of corrections in episodic text. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 52A, 165–183.

Streszczenie. Efekt przedłużonego wpływu dezinformacji (*continued influence effect*, CIE) jest zjawiskiem polegającym na tym, że pewna informacja, mimo że została wycofana i skorygowana, nadal ma wpływ na rozumowanie, wnioskowanie i decyzje. Praktycznym przykładem takiego zjawiska jest szeroko rozpowszechniony mit o tym, że szczepionki powodują występowanie wśród dzieci autyzmu – w związku z tym badania nad CIE mają duże znaczenie z perspektywy praktycznej, ponieważ umożliwiają wgląd w psychologiczne mechanizmy ulegania i rozprzestrzeniania dezinformacji. Celem artykułu jest przedstawienie badań nad efektem przedłużonego wpływu. Omówiono zarówno eksperymentalne metody badania, jak i przedstawiono teorie opisujące poznawcze mechanizmy ulegania dezinformacji. Przytoczono także najważniejsze odkrycia w tej dziedzinie, włączając w to indywidualne i proceduralne wyznaczniki wielkości ulegania dezinformacji oraz techniki zwiększające skuteczność korekt.

Słowa kluczowe: dezinformacja, efekt przedłużonego wpływu, modele mentalne, selektywne wydobywanie, wyznaczniki wielkości ulegania dezinformacji

Data wpłynięcia: 27.05.2021

Data wpłynięcia po poprawkach: 29.11.2021

Data zatwierdzenia tekstu do druku: 1.12.2021