Pourquoi un démenti exact échoue quand même

Tyler Kohler . 4 août 2026 . note 03

Le protocole de 1988

Wilkes et Leatherbarrow ont publié en 1988, dans le Quarterly Journal of Experimental Psychology, l’expérience qui sert encore de paradigme à tout ce domaine. Des participants lisent une suite de messages numérotés sur l’incendie d’un entrepôt ; l’un d’eux signale des pots de peinture à l’huile et des bouteilles de gaz entreposés dans un local, qu’un message ultérieur déclare vide. Interrogés ensuite sur l’origine de la fumée épaisse, les participants continuent de citer la peinture, alors même qu’ils restituent la correction sans erreur quand on la leur demande directement.

La littérature appelle cela l’effet d’influence continue : l’information démentie reste disponible en mémoire et alimente les inférences de gens qui la savent fausse. Hollyn Johnson et Colleen Seifert ont repris ce matériel en 1994 et testé la variante qui intéresse la pratique. Lorsque la correction fournit une autre cause plausible à l’incendie, les références à la peinture diminuent nettement ; elles se maintiennent quand la correction se borne à retirer l’information.

Une découverte qui n’a pas répliqué

Le terme de backfire effect, retour de flamme, vient d’un article de Brendan Nyhan et Jason Reifler paru en 2010 dans Political Behavior. Des participants conservateurs, après avoir lu une correction sur l’absence d’armes de destruction massive en Irak, se déclaraient plus convaincus qu’avant de leur existence. Les conditions comptent pour la suite : quatre études conduites en 2005 et 2006 sur des étudiants de premier cycle, 130 participants pour la première, un effet limité au sous-groupe le plus conservateur et à certains énoncés. Les auteurs demandaient eux-mêmes une réplication hors population étudiante.

Le terme a fait carrière dans la presse, puis dans les manuels de communication de crise, où il a produit une consigne simple : ne jamais répéter l’accusation, même pour la nier. Jusqu’en 2016, elle s’appuyait sur les seules données disponibles. La réplication a échoué à partir de 2017. Kathryn Haglin n’a pas retrouvé, cette année-là, un retour de flamme documenté sur la vaccination antigrippale. En 2019, Thomas Wood et Ethan Porter ont publié cinq expériences, plus de 10'100 participants et 52 affirmations parmi les plus clivantes du débat américain, sans trouver une seule correction capable de déclencher l’effet. L’année suivante, Ullrich Ecker, Stephan Lewandowsky et Matthew Chadwick ont testé la version mémorielle de l’hypothèse, celle qui veut qu’une correction installe la croyance fausse chez un public ignorant la rumeur : 3 expériences, 1'718 participants, résultat globalement négatif, avec des variations entre expériences que les auteurs signalent. Nyhan a écrit en 2021, dans les comptes rendus de l’Académie des sciences américaine, que le retour de flamme n’explique pas la persistance des croyances erronées.

Les ordres de grandeur

En 2017, Man-pui Sally Chan, Christopher Jones, Kathleen Hall Jamieson et Dolores Albarracín ont agrégé 52 études et 6'878 participants. Leur premier chiffre donne l’échelle et circule peu : l’exposition initiale à l’information fausse produit des d situés entre 2,41 et 3,08. Le démenti, lui, obtient des d de 1,14 à 1,33, que les auteurs qualifient d’effets larges, et laisse subsister une persistance comprise entre 0,75 et 1,06. Deux modérateurs ressortent : le détail explicatif de la

correction, corrélé positivement à son efficacité, et les arguments que le public a produits lui-même en faveur de l’information fausse avant qu’on la corrige, qui aggravent la persistance.

La méta-analyse de Nathan Walter et Riva Tukachinsky ne donne pas le même ordre de grandeur. Sur 32 études et 6'527 participants, le résidu d’influence continue s’établit à r égal moins 0,05, avec un p de 0,045, effet minuscule et tout juste significatif. Les deux équipes n’agrègent ni les mêmes études ni les mêmes quantités, et l’écart n’a pas été arbitré dans la littérature.

Une troisième méta-analyse porte sur la vérification journalistique : 30 études, 20'963 participants, effet moyen d égal 0,29, soit un déplacement faible de la moyenne. Le corpus y est fait très majoritairement d’expériences par questionnaire en ligne, avec du matériel réaliste.

Le délai et la source

Briony Swire, Ullrich Ecker et Stephan Lewandowsky ont mesuré la croyance une semaine, puis trois semaines après la correction. Elle remonte vers son niveau antérieur. Le degré de détail explicatif prolonge le bénéfice ; l’avantage d’une affirmation factuelle sur une simple dénégation, net à une semaine, avait disparu à trois semaines. Les participants de plus de 65 ans maintenaient moins bien le résultat de la correction.

Walter et Tukachinsky ont aussi testé des modérateurs, dont quatre intéressent directement la pratique. L’effet correcteur faiblit quand l’information fausse a déjà été répétée avant la correction, et quand elle a été attribuée à une source que le public juge crédible. Il faiblit encore à mesure que s’allonge l’intervalle entre l’exposition et la correction. La configuration la plus favorable mesurée dans ce corpus est celle où la correction émane de la source qui a diffusé l’information. La méta-analyse ne chiffre pas le gain associé à cette configuration, ce qui prive le praticien du moyen de l’arbitrer contre son coût : obtenir d’une rédaction qu’elle corrige elle-même, pièces à l’appui, avant que d’autres supports reprennent. Les études agrégées la mesurent dans des dispositifs où la source est une fiction du protocole.

La croyance et le jugement

En 2020, Nyhan, Porter, Reifler et Wood ont exposé des participants à des vérifications journalistiques réalistes de déclarations faites pendant la campagne présidentielle américaine de 2016. Les croyances factuelles se sont déplacées dans le sens de la vérification, y compris chez les partisans du candidat visé, sans que l’évaluation du candidat suive.

Emily Thorson a proposé en 2016 le terme d’écho de croyance pour désigner ces évaluations négatives qui survivent à une correction acceptée. Jannik Fenger et Martin Vinæs Larsen ont publié en 2026, dans la même revue, une réplication préenregistrée de deux des trois expériences d’origine, même matériel, échantillon plus large. Dans celle où la correction s’est révélée efficace, ils ne trouvent aucun effet attitudinal résiduel. Dans l’autre, la correction n’a pas assez dissipé la croyance initiale pour qu’on puisse séparer les deux effets. Ces mesures portent sur des échelles de sympathie envers un candidat en campagne. Aucune ne porte sur une décision professionnelle, et la transposition reste une hypothèse non testée.

L’audience du démenti

Les expériences citées jusqu’ici supposent toutes que la correction est lue. Andrew Guess, Brendan Nyhan et Jason Reifler ont suivi le trafic web de 2'525 Américains du 7 octobre au 14 novembre 2016. Environ 27,4 % ont visité au moins un site de fausses informations, et les auteurs écrivent

n’avoir presque jamais observé un participant lire la vérification d’une affirmation précise qu’il avait lue. Dans les dossiers que je vois, la correction est publiée sur le site de la personne visée, et sa reprise dépend d’une décision de rédaction sur laquelle l’intéressé n’a aucune prise.

La diffusion des rumeurs est asymétrique, avec une réserve sérieuse. Soroush Vosoughi, Deb Roy et Sinan Aral ont analysé environ 126'000 cascades sur Twitter entre 2006 et 2017 : le centile supérieur des cascades fausses atteignait de 1'000 à 100'000 personnes, quand les cascades vraies dépassaient rarement 1'000 personnes. En 2021, Jonas Juul et Johan Ugander ont montré que ces différences de structure s’effacent dès lors qu’on compare des cascades de taille égale. Ce qui distinguerait les deux tient alors à la contagiosité initiale, les chemins de propagation étant comparables à taille donnée.

Une couche s’est ajoutée depuis 2024 avec les résumés automatiques affichés en tête des résultats de recherche. Le Pew Research Center a suivi la navigation de 900 adultes américains en mars 2025 : 18 % des recherches Google observées produisaient un tel résumé, le clic sur un résultat classique tombait à 8 % des visites concernées contre 15 % sans résumé, et les sources citées dans le résumé étaient cliquées dans 1 % des cas. Ces chiffres portent sur des clics, et aucune étude publiée ne mesure à ce jour ce que devient un démenti à l’intérieur d’un tel résumé.

Les conditions qui améliorent le résultat

Deux éléments reviennent d’une étude à l’autre : le détail explicatif de la correction, qui prédit la durée du bénéfice chez Swire et ses coauteurs, et l’explication causale de remplacement, testée dès 1994 par Johnson et Seifert. La consigne héritée de 2010, qui interdit de répéter l’accusation pour la nier, n’a plus de base empirique depuis les travaux publiés en 2019 et 2020.

Une limite commune à ces protocoles mérite d’être notée pour finir : la lecture de la correction y est garantie par le dispositif expérimental. Les taux d’exposition relevés par Guess et ses coauteurs donnent la mesure de l’écart avec la situation ordinaire.

Wilkes, A. L., et Leatherbarrow, M., « Editing episodic memory following the identification of error », The Quarterly Journal of Experimental Psychology Section A, volume 40, numéro 2, 1988, pages 361 à 387.

Johnson, H. M., et Seifert, C. M., « Sources of the continued influence effect: When misinformation in memory affects later inferences », Journal of Experimental Psychology: Learning, Memory, and Cognition, volume 20, numéro 6, 1994, pages 1420 à 1436.

Nyhan, B., et Reifler, J., « When corrections fail: The persistence of political misperceptions », Political Behavior, volume 32, numéro 2, 2010, pages 303 à 330.

Haglin, K., « The limitations of the backfire effect », Research and Politics, volume 4, numéro 3, 2017.

Wood, T., et Porter, E., « The elusive backfire effect: Mass attitudes’ steadfast factual adherence », Political Behavior, volume 41, 2019, pages 135 à 163.

Ecker, U. K. H., Lewandowsky, S., et Chadwick, M., « Can corrections spread misinformation to new audiences? Testing for the elusive familiarity backfire effect », Cognitive Research: Principles and Implications, volume 5, article 41, 2020.

Nyhan, B., « Why the backfire effect does not explain the durability of political misperceptions », Proceedings of the National Academy of Sciences, volume 118, numéro 15, 2021.

Chan, M. S., Jones, C. R., Hall Jamieson, K., et Albarracín, D., « Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation », Psychological Science, volume 28, numéro 11, 2017, pages 1531 à 1546. Ampleurs rapportées : exposition initiale, d de 2,41 à 3,08 ; démenti, d de 1,14 à 1,33 ; persistance, d de 0,75 à 1,06.

Walter, N., et Tukachinsky, R., « A meta-analytic examination of the continued influence of misinformation in the face of correction », Communication Research, volume 47, numéro 2, 2020, pages 155 à 177. Résidu rapporté : r égal moins 0,05, p égal 0,045.

Walter, N., Cohen, J., Holbert, R. L., et Morag, Y., « Fact-checking: A meta-analysis of what works and for whom », Political Communication, volume 37, numéro 3, 2020, pages 350 à 375.

Swire, B., Ecker, U. K. H., et Lewandowsky, S., « The role of familiarity in correcting inaccurate information », Journal of Experimental Psychology: Learning, Memory, and Cognition, volume 43, numéro 12, 2017, pages 1948 à 1961.

Nyhan, B., Porter, E., Reifler, J., et Wood, T. J., « Taking fact-checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability », Political Behavior, volume 42, numéro 3, 2020, pages 939 à 960.

Thorson, E., « Belief echoes: The persistent effects of corrected misinformation », Political Communication, volume 33, numéro 3, 2016, pages 460 à 480.

Fenger, J., et Larsen, M. V., « Do beliefs echo? On the persistent effects of misinformation after effective corrections », Political Communication, 2026, DOI 10.1080/10584609.2026.2623049.

Guess, A., Nyhan, B., et Reifler, J., « Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign », document de travail soutenu par le Conseil européen de la recherche, diffusé par Dartmouth College, l’université de Princeton et l’université d’Exeter, 2018. Trafic web de 2'525 personnes suivi du 7 octobre au 14 novembre 2016.

Vosoughi, S., Roy, D., et Aral, S., « The spread of true and false news online », Science, volume 359, numéro 6380, 2018, pages 1146 à 1151.

Juul, J. L., et Ugander, J., « Comparing information diffusion mechanisms by matching on cascade size », Proceedings of the National Academy of Sciences, volume 118, 2021, article e2100786118.

Pew Research Center, « Google users are less likely to click on links when an AI summary appears in the results », 22 juillet 2025. Navigation de 900 adultes américains suivie en mars 2025.

Kohler, T. (2026). Pourquoi un démenti exact échoue quand même. Notes, tylerkohler.ch.
https://tylerkohler.ch/notes/pourquoi-un-dementi-exact-echoue/