## B ELECTRONIC SUPPLEMENT

Table 2. Categorization of 172 publications regarding user intervention characteristics and methodological aspects.

| Author | Year | Sample Size | Lab study | Online experiment | Field study | Survey | Interviews | Conceptual | Facebook | Twitter/X | General | Other | Social media posts | Articles | Text | Images | Video | Other | Warning | Correction/debunking | Showing indicators | (Binary) label | Highlighting design | Visibility reduction | Removal | Complicate sharing | Specific visualization | Other | Active | Passive | Neither/unclear | Pre exposure | During | At the moment of sharing | On request | Post exposure | Other | Mis-/Disinformation | Rumors | News credibility | Other | Browser extension/Plugin | Own platform | Game | Other |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Author** | **Year** | **Sample Size** | **Evaluation Type** | | | | | | **Platform** | | | | **Format** | | | | | | **Intervention Design** | | | | | | | | | | **Intera.** | | | **Timing** | | | | | | **Concept** | | | | **Implem.** | | | |
| Agley et al. [2] | 2021 | 1,000 | | ● | | | | | | | ● | | | | | | | ● | | | | | | | | | | ● | ● | | | ● | | | | | | ● | | | | | | ● | |
| Aird et al. [3] | 2018 | 370 | ● | ● | | | | | | | ● | | | | | | | ● | | ● | | | | | | | | | | ● | | | ● | | | | | | | | ● | | | ● | |
| Almaliki [4] | 2019 | 100 | | | | ● | | | | | ● | | ● | | | | | | | | | | | | | | | ● | | ● | | | | | | | ● | ● | | | | | | | ● |
| Amin et al. [6] | 2021 | 38 | | ● | | | | | | | | ● | | | | | | ● | | | | | ● | | | ● | | | ● | | | | ● | ● | | | | ● | | | | | ● | | |
| Andi and Akesson [7] | 2020 | 1,003 | | | | ● | | | | | ● | | ● | | | | | | ● | | | | | | | | | | ● | | | | ● | | | | | ● | | | | | | | ● |
| Ardevol-Abreu et al. [8] | 2020 | N1=31 N2=350 | | | | ● | ● | | | | ● | | ● | | | | | | ● | | | | | ● | | | | | ● | | | | ● | | | | | ● | | | | | | | ● |
| Aslett et al. [9] | 2022 | N1=3,862 N2=3,337 N3=968 | | | ● | ● | | | | | | ● | | | | | | ● | | | | ● | ● | | | | | | | ● | | | ● | | | | | | | ● | | ● | | | |
| Autry and Duarte [11] | 2021 | N1=357 N2=75 | ● | | | | | | | | ● | | | | | | | ● | | ● | | | | | | | | | | ● | | | | | | ● | | ● | | | | | | ● | |
| Axelsson et al. [12] | 2021 | N1=90 N2=119 | | ● | | | | | | | ● | | ● | | | ● | ● | | | | | | | | | | | ● | | | ● | | | | | | ● | ● | | | | | | | ● |
| Ayoub et al. [13] | 2021 | 244 | | ● | | | | | | | ● | | | | ● | | | | | | ● | | | | | | | | | ● | | | ● | | | | | ● | | | | | ● | | |
| Bachmann and Valenzuela [14] | 2023 | 1,472 | | ● | | | | | ● | | | | ● | | | | | | | ● | ● | | | | | | | | | ● | | | ● | | | | | ● | | | | | | | ● |
| Barman and Colan [16] | 2023 | 348 | | ● | | | | | | | ● | | ● | | | | | | ● | ● | | | | | | | | | | ● | | | ● | | | | | ● | | | | | | | ● |
| Barua et al. [17] | 2019 | - | | | | | | ● | | | | ● | | | ● | | | | | | | ● | | | | | | | | | ● | | | | ● | | | ● | | | | | ● | | |
| Bhuiyan et al. [18] | 2021 | N1=430 N2=12 | ● | | ● | | | | | ● | | | ● | | | | | | ● | | | | ● | ● | | | | | | ● | | | ● | | | | | | | ● | | ● | | | |
| Bhuiyan et al. [19] | 2021 | 31 | ● | | | | ● | | | | | ● | | ● | | | | | | | ● | | ● | | | | | | | ● | | | ● | | | | | | | ● | | | | | ● |
| Bhuiyan et al. [20] | 2018 | 16 | | | ● | | | | | ● | | | ● | | | | | | ● | | | | ● | ● | | | | | | ● | | | ● | | | | | | | ● | | ● | | | |

Continued from previous page

| Author | Year | Sample Size | Lab study | Online experiment | Field study | Survey | Interviews | Conceptual | Facebook | Twitter/X | General | Other | Social media posts | Articles | Text | Images | Video | Other | Warning | Correction/debunking | Showing indicators | (Binary) label | Highlighting design | Visibility reduction | Removal | Complicate sharing | Specific visualization | Other | Active | Passive | Neither/unclear | Pre exposure | During | At the moment of sharing | On request | Post exposure | Other | Mis-/Disinformation | Rumors | News credibility | Other | Browser extension/Plugin | Own platform | Game | Other |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Author** | **Year** | **Sample Size** | **Evaluation Type** | | | | | | **Platform** | | | | **Format** | | | | | | **Intervention Design** | | | | | | | | | | **Intera.** | | | **Timing** | | | | | | **Concept** | | | | **Implem.** | | | |
| Bode and Vraga [22] | 2015 | N1=524 N2=500 | | | |•| | |•| | | |•| | | | | | |•| | | | | | | | | |•| | |•| | | | |•| | | | | | |•|
| Bode and Vraga [23] | 2018 | 136 | |•| | | | |•| | | |•| | | | | | |•| | | | | | | | | |•| | |•| | | | |•| | | |•| | | |
| Bozarth et al. [24] | 2023 | 18 | | | | |•| | | | |•|•| | | | | | |•|•|•| | | | | | | | |•| |•| | | | |•| | | | | | |•|
| Brashier et al. [25] | 2021 | 2,683 | |•| | | | | | |•| |•| | | | | | | | |•| | | | | | | |•| |•|•| | |•| |•| | | | | | |•|
| Buczel et al. [28] | 2024 | 337 | |•| | | | | | |•| | |•| | | | |•| | | | | | | | | |•|•| |•| | | |•| |•| | | | | | |•|
| Capraro and Celadin [30] | 2022 | N1=550 N2=558 N3=550 N4=372 | |•| | | | |•| | | |•| | | | | | | | | | | | | | |•| | |•| | |•| | | |•| | | | | | |•|
| Caramancion [31] | 2022 | 327 | |•| | | | | | |•| | | | | | |•| |•|•| | | | | | |•|•| | |•| | | | | |•| | | | | | |•|
| Challenger et al. [32] | 2022 | N1=1,291 N2=2,084 | |•| | | | | | |•| | | | | | |•| |•| | | | | | | | | | |•| | | | | |•| | | |•| | | |•|
| Chen et al. [34] | 2022 | 10 | | | | |•| |•| | | |•| | | | | | | | | | | | | |•| | | |•| | | |•| | |•| | | | | | |•|
| Chen and Tang [35] | 2022 | 415 | |•| | | | | | |•| | | | | | |•| | | | | | | | | |•|•| | |•| | | | | |•| | | | | | |•|
| Chiang et al. [37] | 2022 | 60 | |•| | | | | | |•| | |•| | | | | | |•| | | | | | | | | |•| | | |•| | | | |•| | |•| | |
| Clayton et al. [38] | 2020 | 2,994 | | | |•| | |•| | | |•| | | | | |•| | | | | | | | |•| |•| |•|•| | | | |•| | | | | | |•|
| Craig and Vijaykumar [39] | 2023 | 231 | |•| | | | | | | |•| | |•| | | | |•| | | | | | | |•| |•| |•| | | |•| |•| | | | | | |•|
| Dai [41] | 2021 | 350 | |•| | | | | | |•| | | |•| | | | |•| | | | | | | |•| |•| |•| | | |•| |•| | | | | | |•|
| Dai et al. [40] | 2022 | N1=425 N2=625 | | | |•| | | | |•| | | |•| | | | |•| | | | | | | | | | |•| |•| | | | |•| | | | | | |•|
| Danry et al. [42] | 2020 | 18 |•| | | | | | | | |•| | | | | |•| | | | | | | | | |•| | |•| |•| | | | |•| | | | | | |•|
| Denner et al. [43] | 2023 | 211 | |•| | | | |•| | | |•| | | | | | |•| | | | | | | | | |•| | | | | |•| |•| | | | | | |•|
| Desai and Reimers [44] | 2023 | 365 | |•| | | | | |•| | |•| | | | | | |•| | | | | | | | | |•| | | | | |•| |•| | | | | | |•|
| Dobber et al. [45] | 2023 | 1,054 | |•| | | | | | | |•| | | | |•| |•| | |•| | | | | | |•|•| |•|•| | | | | | | |•| | | |•|
| Domgaard and Park [46] | 2021 | 250 | |•| | | | | | |•| | |•| | | | | | |•| | | | | |•| | |•| |•| | | | | |•| | | | | | |•|
| Drolsbach and Pröllochs [47] | 2023 | 7 | |•| | | | | |•| | |•| | | | | | |•| | | | | | | | | |•| | |•| | | | |•| | | | | | |•|
| Duncan [48] | 2020 | 390 | | | |•| | | | | |•| | |•| | | | | | | | | | | | |•| |•| | |•| | | | | | |•| | | | |•|
| Ecker et al. [49] | 2020 | 2,279 | |•| | | | | | | |•| |•| | | | |•| | | | | | | | | |•| | | | | |•| |•| | | | | | |•|

Continued from previous page

| Author | Year | Sample Size | Lab study | Online experiment | Field study | Survey | Interviews | Conceptual | Facebook | Twitter/X | General | Other | Social media posts | Articles | Text | Images | Video | Other | Warning | Correction/debunking | Showing indicators | (Binary) label | Highlighting design | Visibility reduction | Removal | Complicate sharing | Specific visualization | Other | Active | Passive | Neither/unclear | Pre exposure | During | At the moment of sharing | On request | Post exposure | Other | Mis-/Disinformation | Rumors | News credibility | Other | Browser extension/Plugin | Own platform | Game | Other |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Author** | **Year** | **Sample Size** | **Evaluation Type** | | | | | | **Platform** | | | | **Format** | | | | | | **Intervention Design** | | | | | | | | | | **Intera.** | | | **Timing** | | | | | | **Concept** | | | | **Implem.** | | | |
| Ecker et al. [50] | 2017 | 60 | • | | | | | | | | | • | | • | | | | | | • | | | | | | | | | | • | | | | | | • | | | | | • | | | | • |
| Ecker et al. [52] | 2011 | N1=161 N2=138 | • | | | | | | | | | • | | • | | | | | | • | | | | | | | | | | • | | | | | | • | | | | | • | | | | • |
| Ecker et al. [51] | 2020 | 1718 | | • | | | | | | | • | | • | | | | | | | • | | • | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Feng et al. [55] | 2023 | 595 | | • | | | | | | • | | | | | | • | • | | | | • | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Figl et al. [56] | 2023 | 256 | | • | | | | | • | | | | • | | | | | | • | | | • | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Folkvord et al. [57] | 2022 | 307 | | • | | | | | • | | | | • | | | | | | • | | | | | | | | | • | • | | | • | | | | | | • | | | | | | | • |
| Freeze et al. [59] | 2021 | 434 | | • | | | | | | | | • | | • | | | | | • | | | | | | | | | | | • | | | | | | • | | • | | | | | | | • |
| Furuta and Suzuki [60] | 2021 | - | | | | | | • | | | | • | | | • | | | | | | • | | • | | | | | | | • | | | | | | | • | • | | | | | | | • |
| Gao et al. [62] | 2018 | 122 | | • | | | | | | | • | | • | | | | | | • | | | | | | | | | • | | • | | | • | | | | | • | | • | | | • | | |
| Gesser-Edelsburg et al. [63] | 2018 | 243 | • | • | | | | | • | | | | • | | | | | | | • | | | | | | | | | | • | | | | | | • | | | | | • | | | | • |
| Grady et al. [64] | 2021 | 418 | | | | • | | | | | • | | • | | | | | | • | | | • | | | | | | | • | • | | • | • | | | • | | • | | | | | | | • |
| Grandhi et al. [65] | 2021 | 376 | | | | • | | | | | • | | • | | | | | | | | • | | | | | | | | | | • | | | | | | • | | | • | | | | | • |
| Guess et al. [67] | 2020 | N1=9,190 N2=4,669 N3=6,439 | • | • | | | | | • | | | | • | | | | | | | | • | | | | | | | | | • | | • | | | | | | • | | | | | | | • |
| Guo et al. [69] | 2023 | 28 | | | | | • | | | | | • | | | | | • | | • | • | | | | • | | | | | • | • | | • | • | | | | | • | | | | | | | • |
| Hameleers [70] | 2020 | 1,091 | | • | | | | | | | • | | | • | | | | | • | • | • | | | | | | | | | • | | • | • | | | | | • | | | | | | | • |
| Hameleers et al. [71] | 2020 | 1,404 | | • | | | | | | • | | | • | | | | | | | • | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Hameleers and van der Meer [72] | 2023 | 1,105 | | • | | • | | | | | • | | | | • | | | | • | | • | | | | | | | | • | | | • | | | | | | • | | | | | | | • |
| Hartwig et al. [74] | 2024 | N1=21 N2=18 | | | | • | • | | | | | • | | | | | • | | | | • | | • | | | | | | | • | | | | | • | | | • | | | | | | | • |
| Hartwig et al. [77] | 2024 | N1=44 N2=23 | | | | • | • | | | • | | | • | | | | | | | | • | | • | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Hartwig et al. [76] | 2024 | 20 | | | | | • | | | | | • | | | | | | • | | | • | | • | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Hawa et al. [78] | 2021 | - | | | | | | • | | | | • | | | • | | | | | | | • | | | | | | | | | • | | | | • | | | • | | | | | • | | |
| Heuer and Glassman [80] | 2022 | N1=188 N2=208 | | • | | | | | | | • | | | • | | | | | | | • | | • | | | | | • | | • | | | • | | | | | • | | | | | | | • |
| Horne et al. [81] | 2019 | - | | | | | | • | | | • | | • | | | | | | | | | | | | | | | • | | | • | | • | | | | | • | | | | | | | • |
| Huang and Wang [82] | 2020 | N1=235 N2=235 | | • | | | | | • | | | | • | | | | | | | • | | | | | | | | | | • | | | | | | • | | • | | | | | | | • |
| Irving et al. [83] | 2022 | 129 | | • | | | | | | | • | | | | • | | | | | • | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |

Continued from previous page

| Author | Year | Sample Size | Lab study | Online experiment | Field study | Survey | Interviews | Conceptual | Facebook | Twitter/X | General | Other | Social media posts | Articles | Text | Images | Video | Other | Warning | Correction/debunking | Showing indicators | (Binary) label | Highlighting design | Visibility reduction | Removal | Complicate sharing | Specific visualization | Other | Active | Passive | Neither/unclear | Pre exposure | During | At the moment of sharing | On request | Post exposure | Other | Mis-/Disinformation | Rumors | News credibility | Other | Browser extension/Plugin | Own platform | Game | Other |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Author** | **Year** | **Sample Size** | **Evaluation Type** | | | | | | **Platform** | | | | **Format** | | | | | | **Intervention Design** | | | | | | | | | | **Intera.** | | | **Timing** | | | | | | **Concept** | | | | **Implem.** | | | |
| Jahanbakhsh and Karger [85] | 2024 | 32 | | • | | • | | | | | • | | • | | | | | | | | | • | | | | | • | | | • | | | • | | | | | • | | | | • | | | |
| Jahanbakhsh et al. [89] | 2022 | N1=154 N2=14 | | | • | • | | | | | • | | • | | | | | | | | | • | • | | | | • | | • | | | | • | • | | | | • | | | | | • | | |
| Jahanbakhsh et al. [87] | 2021 | 1,668 | | | | • | | | | | • | | • | | | | | | | | | | | | | • | | | • | | | | | • | | | | • | | | | | | | • |
| Jahanbakhsh et al. [86] | 2023 | 61 | | • | | | | | | • | | | • | | | | | | | | | • | • | | | | | • | | | • | | • | | | | | • | | | | | | | • |
| Jahanbakhsh et al. [88] | 2022 | N1=27 N2=312 | | | • | • | | | | | • | | | • | | | | | | • | | | | | | | • | | | • | | | • | | | | | • | | | | • | | | |
| Jahng et al. [90] | 2021 | 205 | | • | | | | | | | • | | | • | | | | | | • | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Jennings and Stroud [92] | 2021 | N1=1,262 N2=1,586 | | • | | | | | • | | | | • | | | | | | • | | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Jeon et al. [93] | 2024 | N1=6 N2=94 | | • | • | | • | | | | • | | | • | | | | | | | | | | | | | • | | | • | | | | | • | | | | | | • | | • | | |
| Jia et al. [95] | 2022 | 1,677 | | • | | | | | | | • | | • | | | | | | | | | • | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Joshi et al. [96] | 2023 | - | | | | | | • | | | • | | • | | | | | | | | • | | • | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Karduni et al. [97] | 2019 | 5 | | | | | • | | | • | | | • | | | | | | | | • | | | | | | • | | | | • | • | | | • | | | • | | • | | | • | | |
| Kessler and Bachmann [98] | 2022 | 700 | | • | | | | | | | • | | | | | | | • | | • | | | | | | | • | | | • | | | | | | • | | | | | • | | | | • |
| Khivasara et al. [99] | 2020 | - | | | | | | • | | | | • | | | • | | | | | | | • | | | | | | | • | | | | | | • | | | • | | | | • | | | |
| Kim et al. [100] | 2019 | N1=590 N2=299 | | | | • | | | • | | | | • | | | | | | | | | | | | | | | • | | • | | | • | | | | | • | | • | | | | | • |
| Kim et al. [101] | 2019 | - | | | | | | • | | | | • | • | | | | | | | | | | | | | | • | • | | | • | | | | • | | | | | | • | | • | | |
| Kim et al. [102] | 2021 | 92 | • | | | | | | | • | | | • | | | | | | | • | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Kim et al. [103] | 2023 | N1=17 N2=57 N3=49 | • | | | | • | | | • | | | • | | | | | | | | • | | • | • | | • | | | • | • | | • | | | • | | | • | | | | • | | | |
| Kirchner and Reuter [104] | 2020 | N1=1,012 N2=15 N3=1,030 | | | | • | | | • | | | | • | | | | | | • | • | | | | • | | • | | | • | • | | • | • | • | | | | • | | | | | | | • |
| Koch et al. [105] | 2023 | 571 | | • | | | | | • | | | | • | | | | | | • | | • | | | | | | | • | • | • | | | • | | | | | • | | | | | | | • |
| Komendantova et al. [106] | 2021 | N1=103 N2=68 N3=50 | | | | | • | | | | • | | | | | | | • | | | | | | | | | | • | | | • | | | | | | • | • | | | | • | | | |
| Kreps and Kriner [107] | 2022 | 2,000 | | • | | | | | | | • | | | | | | | • | • | • | | • | | • | | | | | • | | | | • | | | | | • | | | | | | | • |
| Lee [108] | 2022 | 171 | | • | | | | | • | | | | • | | | | | | | • | | • | | | | | | | | • | | | | | | • | | • | | | | • | | | • |
| Lee et al. [111] | 2022 | - | | | | | | • | | | | • | | | | | | • | | | • | • | • | | | | • | | | | • | | | | • | | | • | | | | | • | | |
| Lee and Bissell [109] | 2023 | 377 | | • | | | | | • | | | | • | | | | | | • | • | | • | | | | | | | • | • | | • | • | | | | | • | | | | | | | • |

Continued from previous page

| Author | Year | Sample Size | Lab study | Online experiment | Field study | Survey | Interviews | Conceptual | Facebook | Twitter/X | General | Other | Social media posts | Articles | Text | Images | Video | Other | Warning | Correction/debunking | Showing indicators | (Binary) label | Highlighting design | Visibility reduction | Removal | Complicate sharing | Specific visualization | Other | Active | Passive | Neither/unclear | Pre exposure | During | At the moment of sharing | On request | Post exposure | Other | Mis-/Disinformation | Rumors | News credibility | Other | Browser extension/Plugin | Own platform | Game | Other |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Author** | **Year** | **Sample Size** | **Evaluation Type** | | | | | | **Platform** | | | | **Format** | | | | | | **Intervention Design** | | | | | | | | | | **Intera.** | | | **Timing** | | | | | | **Concept** | | | | **Implem.** | | | |
| Lee and Bissell [110] | 2024 | 502 | | • | | | | | • | | | | • | | | | | | | • | | | | | | | | | | • | | | • | | | • | | • | | | | | | | • |
| Liaw et al. [113] | 2013 | ? | | • | | | | | | | | • | | • | | | | | | • | | | • | | | | | | | • | | | • | | | | | • | | | | • | | | |
| Lillie et al. [114] | 2024 | 469 | | • | | | | | | | • | | | • | | | | | | • | | | | | | | | | | • | | | | | | • | | • | | | | | | | • |
| Lim and Perrault [116] | 2023 | 36 | | • | | | | | | | • | | • | | | | | | • | | | • | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Lim and Perrault [115] | 2023 | 200 | | • | | | | | | | • | | • | | | | | | • | | | • | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Liu et al. [117] | 2023 | 859 | | • | | | | | | • | | | • | | | | | | | • | | | | | | | | | | • | | | | | | • | | • | | | | | | | • |
| Lo et al. [118] | 2021 | 89 | | • | | | | | | | • | | | • | | | | | | • | | | | | | | | | | | • | | • | | | | | • | | | | | • | | |
| Lu et al. [119] | 2022 | N1=538 N2=1,098 | | • | | | | | | | • | | | • | | | | | | | | • | | | | | | • | | • | | | • | | | • | | • | | | | | | | • |
| Martel et al. [120] | 2021 | 2,228 | | • | | | | | • | • | | | • | | | | | | | • | | | | | | | | | | • | | | | | | • | | • | | | | | | | • |
| Martino et al. [121] | 2020 | - | | | | | | • | | | | • | | | • | | | | | | • | | • | | | | | | | | • | | | | • | | | | | | • | | • | | |
| Mena [122] | 2020 | 501 | | • | | | | | • | | | | • | | | | | | • | | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Moon et al. [125] | 2022 | 354 | | • | | | | | • | | | | • | | | | | | | • | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Moravec et al. [126] | 2020 | 398 | | • | | | | | • | | | | • | | | | | | • | | | • | | | | | | • | | • | | • | • | | | | | • | | | | | | | • |
| Nekmat [127] | 2020 | 929 | | • | | | | | | | | • | | • | | | | | • | | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Ozturk et al. [129] | 2015 | 259 | | | | • | | | | • | | | • | | | | | | • | • | | | | | | | | | | • | | | • | | | | | | • | | | | | | • |
| Papakyriakopoulos and Goodman [131] | 2022 | - | | | | | | • | | • | | | • | | | | | | • | | | • | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Pareek and Goncalves [132] | 2024 | 96 | | • | | | | | | | • | | | • | | | | | | | | • | | | | | | | | • | | | | | • | | | • | | | | | | | • |
| Park et al. [133] | 2021 | 11,145 | | | | • | | | | | | • | | | • | | | | • | | | | | | | | • | | | • | | | • | | | | | | • | | | | | | • |
| Pasquetto et al. [134] | 2022 | N1=2,805 N2=25 | | • | | | • | | | | | • | | | | | | • | | • | | • | | | | | | • | | • | | | | | | • | | • | | | | | | | • |
| Pennycook et al. [135] | 2020 | N1=5 271 N2=1,568 | | • | | | | | • | | | | • | | | | | | • | | | • | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Pennycook et al. [137] | 2020 | N1=853 N2=856 | | | | • | | | | | • | | • | | | | | | | | | | | | | | | • | • | | | • | | | | | | • | | | | | | | • |
| Pennycook et al. [136] | 2021 | N>5,000 | | | • | • | | | | • | | | • | | | | | | | | | | | | | | | • | | • | | • | | | | • | | • | | | | | | | • |
| Pillai and Fazio [139] | 2023 | 499 | | • | | | | | | | | | | | | | | • | | | | | | | | | | • | | • | | | | | | • | | • | | | | | | | • |

Continued from previous page

| Author | Year | Sample Size | Lab study | Online experiment | Field study | Survey | Interviews | Conceptual | Facebook | Twitter/X | General | Other | Social media posts | Articles | Text | Images | Video | Other | Warning | Correction/debunking | Showing indicators | (Binary) label | Highlighting design | Visibility reduction | Removal | Complicate sharing | Specific visualization | Other | Active | Passive | Neither/unclear | Pre exposure | During | At the moment of sharing | On request | Post exposure | Other | Mis-/Disinformation | Rumors | News credibility | Other | Browser extension/Plugin | Own platform | Game | Other |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Author** | **Year** | **Sample Size** | **Evaluation Type** | | | | | | **Platform** | | | | **Format** | | | | | | **Intervention Design** | | | | | | | | | | **Intera.** | | | **Timing** | | | | | | **Concept** | | | | **Implem.** | | | |
| Pluviano et al. [140] | 2017 | 120 |•| | | | | | | |•| | | |•| | | | |•| | | | | | |•| | |•| | |•| | | | |•| | |•| | | |•|
| Porter and Wood [142] | 2022 | N1=5,000<br>N2=2,000 | |•| | | | |•| | | |•| | | | | |•|•| |•| |•| | | | |•| | | |•| | | | |•| | | | | | |•|
| Porter et al. [141] | 2022 | 5,075 | |•| | | | | | |•| |•| | | | | | |•| | | | | | | | | |•| | |•| | | | |•| | | | | | |•|
| Pourghomi et al. [144] | 2017 | - | | | | | |•|•| | | |•| | |•| | | | | | | | | | |•| | |•| | | | |•| | |•| | | | | | |•|
| Pretus et al. [145] | 2024 | N1=1,709<br>N2=804 | |•| | | | | |•| | |•| | | | | |•| |•| | | | | | |•| |•| | |•| | | | |•| | | | | | |•|
| Prike et al. [146] | 2024 | 415 | |•| | | | | | |•| |•| | | | | | | |•|•| | | | |•| | |•| | |•| | | | |•| | | | | | |•|
| Qian et al. [148] | 2023 | 905 | |•| | | | | | |•| | | | |•| | | | | | | | | | |•| |•|•| |•| | | | | |•| | | | | | |•|
| Rich and Zaragoza [150] | 2020 | N1=134<br>N2=134<br>N3=102 | |•| | | | | | | |•| |•| | | | | |•| | | | | | | | | |•| | | | | |•| |•| | | | | | |•|
| Ruffin et al. [153] | 2022 | N1=113<br>N2=543 | | | |•| | | | |•| | | | |•| | | | |•| |•| | | | | | |•| | |•| | | | | | | |•| | | |•|
| Scharrer et al. [161] | 2022 | 41 | |•| | | | | | |•| | | |•| | | |•| | |•| | | | | | | |•| |•|•| | | | |•| | | | | | |•|
| Safieddine et al. [154] | 2016 | - | | | | | |•| | | |•| | |•|•| | | | | | | | | | |•| | |•| | | | |•| | |•| | | |•|•| | |
| Sakhnini and Chattopadhyay [155] | 2022 | 11 | | | | |•| | | |•| | | | | | |•| |•|•|•| | | | | | | | |•| | | |•| | |•| | | | |•| | |
| Saltz et al. [157] | 2021 | N1=15<br>N2=23 | | | | |•| | | |•| |•| | |•|•| |•| | |•| |•|•| | |•|•|•| | | | | | |•|•| | | | | | |•|
| Sangalang et al. [159] | 2019 | N1=385<br>N2=586 | |•| | | | | | | |•| | |•| | | | |•| | | | | | | | | |•| | | | | |•| |•| | | | | | |•|
| Schaewitz and Kramer [160] | 2020 | 221 | |•| | | | | | | |•| |•| | | | | |•| | | | | | | | | |•| | | | | |•| |•| | | | | | |•|
| Schmid et al. [163] | 2022 | N1=9<br>N2=7 | |•| | |•| | |•| | |•| | | | | | | |•|•| | | | |•| | | |•| | | |•| | |•| | | | |•| | |
| Schmid and Betsch [162] | 2022 | N1=2,444<br>N2=817 | |•| | | | | | |•| | | |•| | | | |•| | | | | | | | |•| | |•| | | |•| |•| | | | | | |•|
| Seo et al. [165] | 2019 | N1=522<br>N2=624 | |•| | | | | | |•| |•| | | | | |•| | | | | | | | | | |•| | |•| | | | |•| | | | | | |•|
| Sharevski and Gover [167] | 2021 | 304 | | | |•| | | |•| | | | | | | |•| | | | | | | | | |•|•|•| | |•| | | | |•| | | | | | |•|
| Sharevski and Zeidieh [168] | 2023 | 29 | | | | |•| | | |•| | | | | |•| |•| | |•| | | | | | | |•| | |•| | | | |•| | | | | | |•|
| Sharevski et al. [166] | 2022 | 337 | |•| | | | | |•| | |•| | | | | |•| | | | | | | | | | |•| | |•| | | | |•| | | | | | |•|
| Sheikh Ali et al. [169] | 2023 | - | | | | | |•| |•| | | | |•| | | | |•|•|•|•| | | | | | |•| | | | |•| | |•| | | | | | |•|

Continued from previous page

| Author | Year | Sample Size | Lab study | Online experiment | Field study | Survey | Interviews | Conceptual | Facebook | Twitter/X | General | Other | Social media posts | Articles | Text | Images | Video | Other | Warning | Correction/debunking | Showing indicators | (Binary) label | Highlighting design | Visibility reduction | Removal | Complicate sharing | Specific visualization | Other | Active | Passive | Neither/unclear | Pre exposure | During | At the moment of sharing | On request | Post exposure | Other | Mis-/Disinformation | Rumors | News credibility | Other | Browser extension/Plugin | Own platform | Game | Other |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Author** | **Year** | **Sample Size** | **Evaluation Type** | | | | | | **Platform** | | | | **Format** | | | | | | **Intervention Design** | | | | | | | | | | **Intera.** | | | **Timing** | | | | | | **Concept** | | | | **Implem.** | | | |
| Sherman et al. [170] | 2021 | N1=24 N2=19 N3=1,456 | | • | | | • | | | | • | | | | | | • | | • | | • | • | • | • | | | • | | | • | | | • | | | | | • | | | | | | | • |
| Smith and Seitz [174] | 2019 | 744 | | • | | | | | • | | | | | | | | | • | | • | | | | | | | | | | • | | | • | | | | | • | | | • | | | | • |
| Song et al. [175] | 2022 | 610 | | • | | | | | • | | | | • | | | | | | | • | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Sotirakou et al. [176] | 2022 | - | | | | | | • | | | • | | | | | | | • | | • | | • | | | | | | • | | | • | | | | • | | | • | | | | | • | | |
| Sullivan [177] | 2019 | N1=625 N2=600 | | • | | | | | • | | | | • | | | | | | | • | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Tanaka and Hirayama [178] | 2019 | 164 | | • | | | | | | • | | | • | | | | | | | • | | | | | | | | | | • | | | | | | • | | | • | | | | | | • |
| Tanaka et al. [179] | 2013 | 87 | | • | | | | | | • | | | • | | | | | | | • | | | | | | | | | | • | | • | | | | • | | | • | | | | | | • |
| Tao et al. [180] | 2023 | 836 | | • | | | | | | | • | | • | | | | | | | • | | | | | | | | | | • | | | | | | • | | • | | | | | | | • |
| Thornhill et al. [182] | 2019 | 20 | | | | • | | | | • | | | | | • | | | | | • | | • | • | | | | | | | • | | | • | | | | | • | | | | | • | | |
| Tseng et al. [183] | 2022 | 210 | | • | | | | | | | • | | | | • | • | • | | | • | | | | | | | | | | • | | | | | | • | | • | | | | | | | • |
| Tsipursky et al. [184] | 2018 | 21 | | | • | | | | • | | | | • | | | | | | | | | | | | | | | • | • | | | • | | | | | | • | | | | | | | • |
| Tulin et al. [185] | 2024 | 752 | | • | | | | | | | • | | | • | | | | | | • | | • | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Tully et al. [186] | 2020 | 610 | | • | | | | | | • | | | • | | | | | | | • | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Tully et al. [187] | 2020 | N1=702 N2=787 | | • | | | | | | • | | | • | | | | | | | | | | | | | | | • | | • | | • | | | | | | • | | | | | | | • |
| van der Meer and Jin [189] | 2020 | 700 | | • | | | | | | | • | | | • | | | | | | • | | | | | | | | | | • | | | | | | • | | • | | | | | | | • |
| van der Meer et al. [188] | 2023 | 1,305 | | • | | | | | | | • | | | • | | | | | • | • | | | | | | | | • | | • | | | | | | • | | • | | | | | | | • |
| Velasco et al. [190] | 2023 | 285 | | | | • | | | | | • | | | | • | | | | | | | • | | | | | | | | | • | | | | • | | | • | | | | • | | | |
| Velez et al. [191] | 2023 | 2,869 | | • | | | | | | | • | | • | | | | | | | • | | | | | | | | | | • | | | | | | • | | • | | | | | | | • |
| von der Weth et al. [192] | 2020 | - | | | | | | • | | • | | | • | | | | | | • | | • | | | | | • | | • | • | | | | | • | | | | • | | | | • | | | |
| Vraga et al. [196] | 2021 | 916 | | • | | | | | | • | | | • | | | | | | | • | | | | | | | | • | | • | | | • | | | | | • | | | | | | | • |
| Vraga and Bode [193] | 2018 | 1,384 | | • | | | | | | • | | | • | | | | | | | • | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Vraga and Bode [197] | 2017 | 271 | | • | | | | | • | • | | | • | | | | | | | • | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Vraga et al. [198] | 2022 | N1=1,207 N2=603 | | • | | | | | | • | | | • | | | | | | | • | | | | | | | | • | | • | | | • | | | | | • | | | | | | | • |

Continued from previous page

| Author | Year | Sample Size | Evaluation Type | | | | | | Platform | | | | Format | | | | | | Intervention Design | | | | | | | | | | Intera. | | | Timing | | | | | | Concept | | | | Implem. | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Author | Year | Sample Size | Lab study | Online experiment | Field study | Survey | Interviews | Conceptual | Facebook | Twitter/X | General | Other | Social media posts | Articles | Text | Images | Video | Other | Warning | Correction/debunking | Showing indicators | (Binary) label | Highlighting design | Visibility reduction | Removal | Complicate sharing | Specific visualization | Other | Active | Passive | Neither/unclear | Pre exposure | During | At the moment of sharing | On request | Post exposure | Other | Mis-/Disinformation | Rumors | News credibility | Other | Browser extension/Plugin | Own platform | Game | Other |
| Vraga et al. [194] | 2021 | 1348 | | • | | | | | • | | | | | | | | • | | | • | | | | | | | | • | | • | | • | • | | | | | • | | | | | | | • |
| Vraga et al. [199] | 2019 | 406 | | • | | | | | • | | | | • | | | | | | | • | | | | | | | | | | • | | | | | | • | | • | | | | | | | • |
| Vraga et al. [195] | 2020 | 1,005 | | • | | | | | | | | • | • | | | | | | | • | | | | | | | | • | • | • | | • | | | | • | | • | | | | | | | • |
| Wahlheim et al. [200] | 2020 | 96 | • | | | | | | | | • | | | | | | | • | | • | | | | | | | | • | | • | | | | | | • | | • | | | | | | | • |
| Waltenberger et al. [201] | 2023 | 9 | | | • | | • | | | | | • | • | | | | | | | | • | • | | | | | | | | • | | | • | | | | | • | | | | • | | | |
| Wang and Huang [204] | 2021 | 271 | | • | | | | | | | | • | • | | | | | | | • | | | | | | | | | | • | | | | | | • | | • | | | | | | | • |
| Wang [202] | 2022 | N1=601 N2=1,060 | | • | | • | | | • | | | • | • | • | | | | | • | • | | | | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Wang et al. [203] | 2022 | 1 | • | | | | | | | | • | | • | | | | | | | | | | | | | | • | | | | • | | | | • | | | | • | | | | • | | |
| Westbrook et al. [207] | 2023 | N1=125 N2=138 N3=251 | | • | | | | | | | • | | | • | | | | | | • | | | | | | | | | | • | | | | | | • | | • | | | | | | | • |
| Wijnker et al. [208] | 2022 | 441 | | • | | | | | | | • | | | | | | | • | • | • | • | | • | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Wood et al. [210] | 2023 | 2,257 | | • | | | | | | | • | | | • | | | | | | • | | | | | | | | | | • | | • | | | | • | | • | | | | | | | • |
| Yong et al. [213] | 2023 | - | | | | | | • | | | • | | | | • | | | | | | • | | • | | | | | | | • | | | • | | | | | • | | | | | | | • |
| Zade et al. [214] | 2023 | 21 | | | | | • | | | • | | | • | | | | | | | | • | | | | | | • | | | • | | | • | | | | | • | | | | | | | • |
| Zhang et al. [216] | 2022 | - | | | | | | • | | | | • | • | | | | | | • | • | | • | | | | | • | | | | • | | • | | | | | • | | | | | | | • |
| Zhao [217] | 2019 | 252 | | • | | | | | • | • | | • | • | | | | | | | • | | | | | | | | | | | • | | | | | | • | • | | | | | | | • |
| Zheng and Ma [218] | 2022 | 222 | | • | | | | | | | • | | | | | | | • | | | | | • | | | | • | | | | • | | • | | | | | • | | | | | | | • |

Table 3. Overview of effects and perceptions of reviewed misinformation interventions (conceptual studies without evaluation were excluded).

| **Source** | **Beneficial Effects** | **Beneficial Perceptions** | **Not effective / counterproductive** |
|---|---|---|---|
| Agley et al. [2] | Exposure to infographics with scientific information slightly increases trust in science compared to exposure to control infographic. | | Exposure to infographics with scientific information does not have direct or indirect effects on COVID-19 preventive behaviors. |
| Aird et al. [3] | Exposure to fact-checks corrects beliefs and affects voters' support when corrections outnumber affirmations compared to other correction ratios and for both sides of the political spectrum ($\eta^2$=0.13 (fact checks); $\eta^2$=0.01 (myth:fact ratio). | | |
| Almaliki [4] | | Users perceive interventions with gamification elements useful but preferences for elements vary. | |
| Amin et al. [6] | Interventions with Visual Selective Attention System can increase attentive behavior of COVID-19 misinformation sharing compared to pre-intervention (D-Scores similar to Cohen's d: Highest number of participants in category 'Neutral/ No Preference' (D-score=-0.15 to D-score=0.15) | | |
| Andi and Akesson [7] | Social norm-based nudge decreases misinformation sharing behavior compared to non-application. | | |
| Ardevol-Abreu et al. [8] | | | Warning labels to assess credibility are not regarded as central assessment measures by users. |
| Aslett et al. [9] | | | Providing dynamic, in-feed source reliability labels do not significantly improve news diet quality or reduce misperceptions (<0.08 change in SD of the pre-treatment measure). |
| Autry and Duarte [11] | | | Negated corrections and replacements lead to increased belief in misinformation for cases with no previous exposure to the target concept, relative to cases with exposure and cases with no treatment ($\eta^2$=0.22 (main effect of exposure); $\eta^2$=0.23(main effect of correction); $\eta^2$=0.18 (interaction between exposure and correction)). |
| Axelsson et al. [12] | Observational learning and feedback as intervention tools increase user performance of credibility assessment compared to the non-treatment control group ($\eta^2$=0.043) | | |
| Ayoub et al. [13] | Additional employment of SHapley Additive exPlanations (SHAP) in NLP misinformation detection model and SHAP combined with source and evidence information increases user trust in misinformation detection compared to presenting output text only. | | |
| Bachmann and Valenzuela [14] | Fact-checks are similarly effective at reducing people's misperceptions across message formats (transparency elements, arousing visuals) (d=0.51 (Study 1) and d=0.38 (Study 2)) | | Compared to control groups without intervention, users exposed to political fact-checks trust news less and perceive the media as more biased, especially after reading corrections debunking pro-attitudinal misinformation. |
| Barman and Colan [16] | Warning flags with and without explanation text from fact-checking websites reduce perceived accuracy of misinformation and intent to share. Explanatory texts could enhance the trustworthiness of the intervention. | | |
| Bhuiyan et al. [18] | Credibility nudges as browser extension improve user's skills to distinguish news tweets' credibility compared to control group (d=0.296) | | |
| Bhuiyan et al. [19] | | Transparency cues (source and message credibility) on news websites increase consumer trust. | |

Continued from previous page

| Source | Beneficial Effects | Beneficial Perceptions | Not effective / counterproductive |
|---|---|---|---|
| Bhuiyan et al. [20] | Attention and reflection nudges enhance users' credibility assessment compared to control group | Attention and reflection nudges enhance users' credibility assessment (reread and rethink news; use external sources; actively participate in assessment) compared to control group. | |
| Bode and Vraga [22] | Exposure to corrective information decreases user misperceptions compared to pre-treatment and to the control group ($\eta^2$=0.052) | | |
| Bode and Vraga [23] | Interventions with algorithmic or social corrections are equally effective in health misinformation corrections compared to control conditions without intervention for high and low conspiracy belief individuals ($\eta^2$=0.046 (interventions overall); $\eta^2$=0.016 (comparison between algorithmic and social correction)) | | |
| Bozarth et al. [24] | | Almost half of participants (moderators on Reddit) preferred cues over labels from expert fact-checkers as they can help discern user intent. A quarter distrusts professional fact-checkers. | |
| Brashier et al. [25] | Debunking measures have a stronger long-term impact on users' fact-checking memory than prebunking, labeling, or no measures. | | |
| Buczel et al. [28] | Warning before misinformation reduces reliance on it in short-term in comparison to no warning. Warning after misinformation had no effect ($\eta^2$=0.05 (forwarning vs. retraction only)) | | Reliance on misinformation increased for over 7 days although the memory of retraction continued. |
| Capraro and Celadin [30] | Accuracy endorsement prompt nudge reduces fake news sharing but also increases sharing of real news compared to simple fake alert and no-nudge (f=0.129 (two nudges); f=0.125 (two nudges, different UI); f=0.129 (comparison between endorsing accuracy condition and accuracy salience condition)) | | |
| Caramancion [31] | | | Preventive infographics have trivial to no effect on social media users |
| Chiang et al. [37] | AI news source credibility system positively affects users' information assessment and attitude towards media literacy learning. | | |
| Challenger et al. [32] | Myth-busting formats, question-answer formats and fact-myth formats are more effective interventions than fact-only formats and control baseline in reducing COVID-19 misinformation agreement ratings. | | |
| Chen and Tang [35] | Intervention with narrative fear appeal messages are effective in promoting health experts to correct online health misinformation for the public. | | |
| Chen et al. [34] | Correct assessment of misinformation overall improved by VisualBubble. Participants became more willing to make assessments and more critical (effect sizes: Topic Filter: large (d=0.98 and 0.98); Opinion Filter: negligible (d=0.00) and medium (d=0.79); Source Filter: large (d=1.11 and 1.01)) | | Showed tendency to become over-skeptical |
| Clayton et al. [38] | Intervention with a general warning about misleading articles reduce the perceived accuracy of false headlines relative to a no-warning condition and 'rated false' tag is more effective than 'disputed' tag.(d=0.08 (general warning before seeing headlines); d=0.26 ('disputed' tag); d=0.38 ('rated false' tag)) | | |
| Craig and Vijaykumar [39] | Corrective infographic improved rating of misinformation as untruthful and reduced reported willingness to share it. Debunking may be short-lived if followed by misinformation. Effect can be maintained in presence of further corrective information (e.g., $\eta^2$=0.150, 0.109 and 0.079) | | |
| Dai [41] | Timing of misinformation correction interventions (pre/post exposure) and addition of coherence message (debiasing/no addition) impacts effectiveness ($\eta^2$=0.087 (post exposure); $\eta^2$=0.047 (debiasing message); $\eta^2$=0.163 (time lapse)) | | |

Continued from previous page

| Source | Beneficial Effects | Beneficial Perceptions | Not effective / counterproductive |
|---|---|---|---|
| Dai et al. [40] | | Most participants indicate that counterfactual explanations can accurately explain why a piece of news is fake and results suggest that the approach generates the most helpful explanations compared to state-of-the-art methods (human evaluation based on survey with young, well-educated participants). | |
| Danry et al. [42] | | Wearable AI system with explainable feedback enhances rationality in evaluating information in comparison to non-explainable AI and control group | |
| Denner et al. [43] | A single correction and repeated corrections significantly increased organizational trust compared with no correction | | Small negative effect of perceived persuasive intent on organizational trust after repeated corrections. |
| Desai and Reimers [44] | | | No evidence that corrections explaining the reason the misinformation was presented were more effective than a correction not accompanied by explanation |
| Dobber et al. [45] | Red and orange traffic light labels placed concurrently with in contrast to prior to the start of a political advertisement significantly affect credibility perception. Direct-to-consumer labels can be effective but it depends on timing and position. | | |
| Domgaard and Park [46] | Interventions with info graphs increase user ability to identify vaccine-related misinformation compared to text-only intervention and no intervention. | | |
| Drolsbach and Pröllochs [47] | Community fact-checked misinformation is less viral and receives fewer retweets than non-misleading posts. | | |
| Duncan [48] | Credibility labels are effective on news validation when ideological perspective of the user match the ideology of the news brand but also in cases where they do not match. | | |
| Ecker et al. [49] | Corrections are generally effective at influencing inferential reasoning but narrative corrections are not more effective than non-narrative | | |
| Ecker et al. [50] | Corrections are more effective when they explicitly repeat the myth compared to corrections that do not repeat the misinformation ($\eta^2$=0.04 (memory); $\eta^2$=0.27 (inferential reasoning)) | | |
| Ecker et al. [52] | Strong corrections and cognitive load interventions, measured in different degrees of interventions or misinformation strength, can reduce (but never fully) the continued influence effect of strong misinformation, but even strong interventions are less effective on weak misinformation. ($\eta^2$=0.05 (strength of misinformation); $\eta^2$=0.41 (strength of correction); $\eta^2$=0.04 (strength of cognitive load on misinformation); $\eta^2$=0.07 (strength of cognitive load on correction);) | | |
| Ecker et al. [51] | Misinformation corrections do not lead to familiarity backfire effects but instead lead to corrective effect in both, audiences unfamiliar to a misinformation and audiences familiar to the topic (i.a., $\eta^2$=0.024 (false claim inference across all conditions: no-exposure/fact-check with and without cognitive load); $\eta^2$=0.004 (fact check condition without cognitive load)) | | |
| Feng et al. [55] | Provenance has effect on credibility perception. Helped correct truth judgments towards deceptive media (qualitatively measured) | | Over-corrected in some cases and shifted away from truth in some non-deceptive media |
| Figl et al. [56] | All evaluated flags lead to reduced perceived credibility. The semantic priming effect of different warning symbols (e.g., stop symbol associated with stopping behavior) makes a difference. Stronger warnings may be required on smartphones than on PCs. | | |

Continued from previous page

| Source | Beneficial Effects | Beneficial Perceptions | Not effective / counterproductive |
|---|---|---|---|
| Folkvord et al. [57] | Interventions with source information positively affect the critical news evaluation compared to a control group with no intervention (e.g., $\eta^2$=0.05 for vaccination misinformation) | | Inclusion of a protective warning message does not significantly affect critical evaluation (e.g., $\eta^2$<0.001 for vaccination and health insurance misinformation). |
| Freeze et al. [59] | General misinformation warnings which also contain invalid instances, in contrast to valid-only instances and control with no intervention, lead to a discarding of authentic information and to increased memory uncertainty. | | |
| Gao et al. [62] | | | Stance labels on political ideologies intensify readers' selective exposure (tendency to look for agreeable opinions), and lower the perception of extremeness and criticality of misinformation. Credibility labels only have a limited effect on reducing selective exposure and misinformation identification. |
| Gesser-Edelsburg et al. [63] | Corrections of misinformation from health organizations are more effective for pro-vaccination as well as for vaccination-hesitant individuals when communication addresses full, transparent information and emotional aspects compared to 'common' one-dimensional, partial responses. | Additional qualitative analysis reinforces quantitative findings. | |
| Grady et al. [64] | Misinformation warnings for political news are effective in short-term to correct beliefs and eliminate partisan bias but in long-term corrected beliefs weaken and biases return. | | |
| Grandhi et al. [65] | | Users perceive trustworthiness indicators as useful for reducing uncertainty and for providing guidance on content interaction. | |
| Guess et al. [67] | Digital media literacy interventions increase user ability to discern between correct information and misinformation compared to control group without intervention (d=0.2 (US-based study); d=0.11 (India-based study)) | | |
| Guo et al. [69] | Specific contextual warnings for video-sharing platforms can alert users to be vigilant and are influenced by explicitness and risk level. In terms of accuracy judgment the interstitial warning and specific contextual warning were both considered effective. | | |
| Hameleers [70] | A combination of media literacy- and fact-checking interventions are most effective in lowering perceived accuracy of political misinformation, compared to each intervention separately and control group without intervention. | | |
| Hameleers and van der Meer [72] | General rather than issue-specific warnings about misinformation are more effective for participants with higher level of trust in the media. | | The prebunking exposure to different warning interventions did not influence the truth rating of factually accurate information or misinformation. Observed negative spillover effects of prebunking warnings on truth rating of accurate information. |
| Hameleers et al. [71] | Multimodality (text-plus-visual) impacts credibility of disinformation but also of fact-checking interventions compared to disinformation and intervention with text-only and compared to control without intervention. | | Source type (ordinary citizen, news agency) does not influence credibility level |
| Hartwig et al. [74] | In several instances, participants changed or consolidated their assessment of the information presented with the help of the indicators. | Participants found the indicators useful for practice and as a reminder to be more able to identify disinformation on their own in the future, without app support. | Adolescents tended to blindly in the intervention. |
| Hartwig et al. [77] | When topical, formal, and rhetorical indicators are presented with tweets, they improve users' perception and evaluation. | Approach is perceived as useful overall within the context of COVID-19 and the Russian war against Ukraine. | |

Continued from previous page

| **Source** | **Beneficial Effects** | **Beneficial Perceptions** | **Not effective / counterproductive** |
|---|---|---|---|
| Hartwig et al. [76] | | When assessing the comprehensibility and perceived usefulness of features to assess a voice message's credibility, it received a mostly positive feedback especially on features that refer to the content itself. | |
| Heuer and Glassman [80] | Checklist with source labels is significantly better in influencing participants' performance on correct article ratings for the better. | Checklist that provides source labels was considered most helpful. The interactive checklist is perceived as more helpful than the written checklist. | |
| Horne et al. [81] | Soft information nudging/trust nudging has potential benefit of moving even extreme or conspiracy news consumers towards higher quality information (based on simulations) | | |
| Huang and Wang [82] | Misinformation belief is impacted by the message format (narrative/non-narrative) and correction mechanism (social/algorithmic correction) ($\eta^2$=0.03 and 0.04 (message format);) | | |
| Irving et al. [83] | Correction reduces number of references to misinformation (medium-to-large effect size) and was remembered and recalled ($\delta$=0.64, 95% BCI [0.28, 0.99] (medium-to-large)) | | |
| Jahanbakhsh and Karger [85] | | It helped them think about the news in a more analytical way or gauge their trust in a source. They liked being interactive with the news content and the ability to call out content they found biased or misleading. | Assessing took extra time and effort. Sometimes they found it hard to assess a piece of content. They want to think for themselves, unassisted by anyone. |
| Jahanbakhsh et al. [89] | Lightweight nudging interventions (checkboxes, checklists, free-text rationales) which provide accuracy assessment and rationale reduce misinformation sharing (but also sharing overall). | | |
| Jahanbakhsh et al. [87] | Users perceive incorporation of three new user affordances into social media as useful tools to independent, user-friendly misinformation combat. | Qualitative examples reinforce quantitative findings. | |
| Jahanbakhsh et al. [86] | Personalized AI impacts users' judgment and grows larger over time, but is reduced when users provide reasoning for their assessment (e.g., exp($\beta$)=1.60 for condition whether AI's prediction had a statistically significant effect on user agreeing with AI) | | |
| Jahanbakhsh et al. [88] | Users perceived value in browser extension that allows to change headlines and used it to make various changes. In follow-up study: substantial number of alternative headlines were preferred especially if bias was removed or deceptions were corrected. | | |
| Jahng et al. [90] | Discounting cues ('fake news' labels) in online comments negatively impact users' ability of veracity evaluation and increase need to authenticate information compared to control group without exposure to discounting cues. (i.a., $\eta^2$=0.041 (evaluation ability); $\eta^2$=0.057 (need to authenticate))) | | |
| Jennings and Stroud [92] | Partisan affiliations impact likeliness to belief in misinformation, particularly about opposing parties (i.a., $\eta^2$=0.13 (user partisanship (P) and party-affiliation of misinformation target(M)); $\eta^2$=0.01 (P, M and fact-check condition (F)) | | Overall, independent from partisan affiliation, fact check interventions do not improve information evaluation compared to cases without intervention. |
| Jeon et al. [93] | Both the quantitative and qualitative results confirmed that HearHere has an impact on mitigating political polarization and broadening one's perspectives on news consumption. | | |
| Jia et al. [95] | Interventions with misinformation labels (algorithm, community, third-party fact-checker, and no label) reduce credibility of misinformation for liberal users independent of post-ideology while only algorithm labels are effective in reducing ideology-consistent misinformation for conservative users (and all label types for opposing-ideology misinformation). | | |

Continued from previous page

| Source | Beneficial Effects | Beneficial Perceptions | Not effective / counterproductive |
|---|---|---|---|
| Karduni et al. [97] | | Visual analytic systems are a helpful tool to support the investigation of misinformation on social media and to enhance traditional (media literacy education) strategies. | |
| Karduni et al. [97] | Corrections of health-related misinformation with additional use of images is more effective in correcting myth belief than without images ($\eta^2$=0.117) | | Image type (machine-technical image, expert image, diagram) does not influence persuasive effect. |
| Kim et al. [100] | Source rating mechanisms are effective interventions to correct users beliefs, whereby expert rating and user article rating are more effective than user source rating. Low ratings and no-ratings have a disproportional stronger effect on user skepticism than high ratings on user trust. | | |
| Kim et al. [101] | Controversy score that provides additional information of opinions on topics and encourages further exploration can be a more effective tool to combat myth belief than approaches that seek to correct or standardize news opinions. | | |
| Kim et al. [103] | | | No single strategy ((1) hiding content, allowing for explanations, and option to toggle view, (2) including an engagement option with the correction that allows for indicator details, (3) Placing agent next to share button that asks for accuracy and reasoning and presents statistics) was superior over the control. Study highlights necessity of transparency and clarity about intervention's logic and concerns about repeated exposure to misinformation and lack of user engagement. |
| Kim et al. [102] | Humorous interventions increase user attention to relevant corrections of misinformation, but non-humorous interventions outperform humorous interventions via higher credibility ratings. ($\eta^2$=0.19 humor) | | |
| Kirchner and Reuter [104] | Warning-based interventions significantly effect perceived news accuracy but explanation-based approaches are most effective. | Warning-based interventions (with additional explanations) are more effective in correcting user beliefs than less transparent methods such as reduced post size and fact-checks in related articles. | |
| Komendantova et al. [106] | | Stakeholders (journalists/fact-checkers, policymakers, citizens) require design tools for mitigating misinformation and prioritise information regarding actors behind misinformation and tracing the life cycle of misinforming posts. The most valued features across groups relate to timing and flow of misinformation. | |
| Koch et al. [105] | Warning labels reduced perceived credibility and lowered self-reported likelihood to amplify fake news (rather small effect). | | Removing social endorsement cues (e.g., engagement counts) did not have an effect. Did not find a positive effect of warning labels on users' likelihood to elaborate on the fake news post. |
| Kreps and Kriner [107] | Compared to no intervention, 'false' tags only have a small effect on users' accuracy assessments while journalistic fact-checks are more effective in reducing misperceptions as well as sharing (independent of partisanship). | | |
| Lee and Bissell [110] | | | Repeated exposure of myths within corrective information increased perceived familiarity about misinformation and increased misinformation credibility (partial $\eta^2$=.02 (effect of correction types on misinformation familiarity)) |

Continued from previous page

| **Source** | **Beneficial Effects** | **Beneficial Perceptions** | **Not effective / counterproductive** |
|---|---|---|---|
| Lee [108] | Web add-on corrections generally decrease the belief in misinformation compared to no correction. For those who are motivated to use social media for specifically for social interaction, narrative corrections are most effective, compared to web add-on's and no corrections ($\eta^2$=0.025 (narrative correction for social interaction-motivated users)) | | Amongst users in general, narrative corrections are not more effective than web add-on corrections or no corrections. |
| Lee and Bissell [109] | Both commenting and AI fact-checking labels were effective at promoting positive attitudes toward vaccination compared to no intervention. Commenting intervention emerged as promising for suburban participants and the AI intervention was pronounced for urban populations ($\eta^2$=.03 (for difference in attitudes between three experimental groups)) | | Neither of the interventions showed salient effects with the rural population. |
| Liaw et al. [113] | | The proposed system utilizes crowd-sourced corrections, such as in-line commentary and corrections which are ranked by the user to enhance comprehension of news. | |
| Lillie et al. [114] | The narrative corrective directly reduced misinformation belief compared with a didactic corrective and a no-correction control. | | |
| Lim and Perrault [116] | Post engagement was generally dampened by the presence of warning labels. | | Participants were more likely to share congruent posts, with or without labels, suggesting the need for other interventions to address political polarization effects. |
| Lim and Perrault [115] | The intent to comment and share was significantly lower for posts with a generic warning label than unlabeled posts. The knowledge, source, and propagation labels encouraged sharing instead. Partisanship effects were observed across the labels (partial $\eta^2$=0.016 for effect of warning labels on sharing intentions and 0.0077 on commenting intention) | | |
| Liu et al. [117] | No differences in effectiveness across fact-checking sources (professional fact-checkers, mainstream news outlets, social media platforms, AI, crowd-sourcing; $\eta^2$=0.01) but sources perceived as more credible are more effective | | |
| Lo et al. [118] | | Indicates effectiveness of an fake news intervention module that co-works with a news recommendation system and guides users towards verified news. | |
| Lu et al. [119] | | | AI label nudges people into aligning their veracity belief in the news with the AI model's prediction regardless of its correctness compared to a control group (Control vs. AI-before: d= 0.17; Control vs. AI-after: d=0.15) |
| Martel et al. [120] | | | Hedging corrections or providing increased explanatory depth in corrections of misinformation had no impact on engagement with corrective messages on social media. |
| Martino et al. [121] | | The Prta system raises awareness about the use of propaganda techniques in the news, promoting media literacy and critical thinking. | |
| Mena [122] | A warning label was effective in reducing the intention of a user to share misinformation on Facebook compared to a user who did not see the warning. (d=0.36) | | |
| Moon et al. [125] | AI and user consensus (vs. human experts) source labels reduced partisan-based motivated reasoning in assessing fact-checking message credibility ($\eta^2$=0.0018 for pattern of motivated reasoning varied by fact-checking sources) | | |
| Moravec et al. [126] | System 1 (automatic cognition) and System 2 (deliberate cognition) interventions both were effective and intervention combining both was twice as effective. | | |

Continued from previous page

| Source | Beneficial Effects | Beneficial Perceptions | Not effective / counterproductive |
|---|---|---|---|
| Nekmat [127] | A fact-check alert was effective in reducing likelihood of sharing misinformation compared to non-exposure. | | |
| Ozturk et al. [129] | A textual counter presented to a rumor was effective in decreasing the likelihood of sharing a tweet compared to the rumor only and rumor with information condition. | | |
| Papakyriakopoulos and Goodman [131] | Textual overlap in labels reduces user interactions and stronger rebuttals reduced toxicity in comments. | | Label placement did not change propensity of users to share and engage with labeled content but falsity of content did. |
| Pareek and Goncalves [132] | Credibility disputes raised by one's co-partisans significantly reduced belief in misinformation, irrespective of one's relationship closeness with the peer. A peer's knowledgeability may be more potent than trustworthiness in causing belief change, and trust can sometimes manifest even in the credibility judgement of distant peers, when perceived to have expertise or a fact-checking tendency. | | |
| Park et al. [133] | When opposite fact-checking labels are shown, users who initially disapprove of a claim are less likely to change their views than those who initially approve of the same claim. | User interviews revealed that users are more likely to share claims with a Divided Evidence label than those with a Lack of Evidence label. | |
| Pasquetto et al. [134] | Audio files on WhatsApp were found to be more effective than text or video-based sources in correcting beliefs about misinformation and they were shared more frequently when communicated by someone close to the user. | | |
| Pennycook et al. [135] | Warnings were effective in a modest reduction in perceived accuracy of false headlines, particularly for politically concordant headlines, relative to a control condition. | | The presence of warnings caused untagged headlines to be seen as more accurate than in the control, even if they were false. |
| Pennycook et al. [137] | Simple accuracy reminders before sharing information on social media are effective in increasing truth discernment in participants' sharing intentions compared to a control group (d=0.142) | | |
| Pennycook et al. [136] | Shifting the attention of the users on the accuracy of information can encourage them to share higher quality news (e.g., Pearson's r=0.71/0.67/0.61) | | |
| Pillai and Fazio [139] | Participants were less likely to share false headlines in the explain prompt condition compared to control group (exceeded the necessary number of participants according to a priori power analysis; $\eta^2$=.03) | | |
| Pluviano et al. [140] | | | Displaying a myth about vaccines causing autism alongside a factual correction resulted in an increase in belief in the myth over a 7 day time period (partial $\eta^2$=0.175) |
| Porter et al. [141] | Corrections eliminate effects of misinformation on beliefs about vaccine. Effect is robust to formatting changes in the presentation of corrections. Corrections without any formatting modifications are effective at reducing false beliefs with formatting variations playing a very minor role (fact-checks increase accuracy by 0.41 scale points on a four-point scale regardless of formatting; modifications to formatting increase accuracy only by 0.03 points.) | | |
| Porter and Wood [142] | Fact-checks are effective in increasing factual accuracy on realistic simulations of social media platforms (Study 1 Correction Effect d=0.55; Study 2 d=0.79) | | |
| Pretus et al. [145] | Adding a misleading count next to the like count reduced participants' reported likelihood to share inaccurate information by 25% compared to control condition. It was five times more effective as an accuracy nudge (misleading count compared to no intervention: d=0.20; misleading count compared to accuracy nudge: d=0.13). | | |

Continued from previous page

| Source | Beneficial Effects | Beneficial Perceptions | Not effective / counterproductive |
|---|---|---|---|
| Prike and Ecker [147] | The social-norm intervention reduced belief in false claims and improved discrimination between true and false claims. It also had some positive impact on social media engagement. Credibility badges led to greater belief in true claims, lower belief in false claims, and improved discrimination. The credibility-badge intervention also had robust positive impacts on social media engagement, leading to increased flagging and decreased liking and sharing of false posts. Credibility badges and social norms can be effective interventions for counteracting online misinformation. Credibility badges were associated with larger effect sizes and more consistent results across the alternative analysis specifications (partial $\eta^2$=0.09 (credibility badge) and 0.01 for social norm ). | | |
| Qian et al. [148] | Active interventions significantly increased intention of using reverse image search tools compared to passive interventions and a control group. | | Neither passive nor active interventions had an effect on credibility judgment or misinformation discernment. |
| Rich and Zaragoza [150] | | | When correcting misinformation, there was no evidence that the time of correction mattered for the efficacy of the correction and the participants corrected beliefs were not durable (durability of corrected belief $\eta^2$=0.43; time of correction $\eta^2$=0.02) |
| Ruffin et al. [153] | Simply highlighting and explaining manipulation in photos was not always effective but when it was, it did help make users less agreeing with intended messages (e.g., $\beta$=-0.58 of linear regression model for explaining the manipulation versus seeing the original image). | | Intervention was not always effective. Explanation had negative effect on feeling/sentiment toward the subject/image |
| Sakhnini and Chattopadhyay [155] | | Fact-checking apps should be sensitive to age-related, personal, and political biases | |
| Saltz et al. [157] | | Findings suggest strong emotional reactions to misinformation labels in general, which are perceived as overly paternalistic, biased, and punitive. | |
| Sangalang et al. [159] | Narrative correctives (with or without emotional ending) can effectively reduce misinformation beliefs, while emotional corrective endings are better at correcting attitudes. | | |
| Schaewitz and Kramer [160] | Detailed corrections presented alongside disinformation are more effective in better remembering facts compared to simple corrections ($\eta^2$=0.02) | | The influence of detailed corrections on personal beliefs regarding the topic of the disinformation is counterproductive as more details in the correction seem to raise readers' concerns when corrections are presented together with the disinformation. |
| Scharrer et al. [161] | Warnings on top of a scientific message made laypeople hesitant about uncritically and confidently accepting the message as true. Participants agreed less with the claims and deemed the text to be less credible than without the warning ($\eta^2$=0.48) | | Warnings cannot reduce or prevent boost in persuasiveness of easily understandable misinformation. |
| Schmid et al. [163] | | A web app based on Social Network Analysis could effectively provide an overview of potentially misleading vs. non-misleading content on Twitter, which can be explored by users and enable foundational learning. | |
| Schmid and Betsch [162] | Text-based refutations effectively reduced belief in misinformation and immunized participants against impact in short-time (final power of 94.5% was reached to detect a small effect size. Credibility judgment after 2 months was slightly lower (d=0.04)) | Unintended effects: lacking effect on intentions, backfire-effects among religious groups, biased judgments when omitting information about vaccine side effects | |
| Seo et al. [165] | Machine-Learning-Graph warning, indicating Source Reliability, Content Truthfulness and Picture/Video Truthfulness, was effective increased participants' sensitivity in differentiating fake from real news. ($\eta^2$=0.018) | | |
| Sharevski and Zeidieh [168] | | | Warning labels as visual frictions are not accessible for low vision or blind users. |

Continued from previous page

| Source | Beneficial Effects | Beneficial Perceptions | Not effective / counterproductive |
|---|---|---|---|
| Sharevski et al. [166] | | SPAM warning tags are promising and increase trust in soft moderation. Text-only variant tells participants more of what is going on and a text-and-flag variant gives more specifics and is tougher to refute as a large visual cue. Warning tag with improbable interpretation of facts (FFS) gave convincing options for users to pick why the context is fitting to the misinformation tweet. Left- and right-leaning participants positively rated the intervention. | |
| Sharevski and Gover [167] | | | The utterance of a warning cover before a Tweet containing valid information about COVID-19 vaccines by Alexa will not reduce the perceived accuracy of the spoken back Tweet's content relative to a no warning cover condition (d=0.018) |
| Sherman et al. [170] | A combination of expert and user insights is effective in defining interpretable warnings and design guidelines for communicating the provenance of video content to end-users. | | results raise concerns around the potential for users to overgeneralize misinformation warnings regarding video or text information |
| Smith and Seitz [174] | Corrective mock Facebook news feeds were effective in reducing belief in neuroscience myths when shown immediately after the misinformation for those who held incorrect beliefs at pretest. | | If participants held correct beliefs at pretest, a single exposure to misinformation (even when immediately corrected) was enough to have a negative impact on their beliefs. |
| Song et al. [175] | Image-only modality triggered significantly lower levels of message elaboration and heightened message credibility and increased engagement intentions (effect of evidence type on self-reported message elaboration: $\eta^2$=0.01. Effect of presentation mode on message elaboration: $\eta^2$=0.02) | Presence of statistical evidence in assertions reduced message elaboration and effects of message in correcting misperceptions, decreased perceived message credibility and lowered intentions to further engange with and disseminate the corrective message. | |
| Sullivan [177] | | | Libraries were not effective in correcting misconceptions about the flu vaccine through comments on social media. |
| Tanaka and Hirayama [178] | Objective countermessages reduced belief in rumors and subjective countermessages strengthened false beliefs (e.g., $\eta^2$=0.02. Post-hoc power analysis revealed adequate G*Power >0.80 at medium to large effect size levels). | | Subjective countermessages even strengthened false beliefs |
| Tanaka et al. [179] | Displaying criticism of false information prior to rumors during a disaster response is effective in increasing proportion of responses aimed at stopping the spread of rumors compared to displaying the criticism after the rumor. | | |
| Tao et al. [180] | All three types of corrections improved belief accuracy. Corrections incorporating hope appeals showed enhanced effectiveness when threat information was present in comparison to absent hope appeals (Power analysis reveals study can detect small effect sizes (f=0.11) with power of 80%. Hope appeal when threat was present versus absent: $\eta^2$=0.01) | | |
| Thornhill et al. [182] | BalancedView, a proof-of-concept that shows news stories relevant to a tweet suggests that nudging users by providing context information may change the behavior of them towards that of informed news readers. | | |
| Tseng et al. [183] | Corrective information in the form of text, images, or videos is effective in reducing participants' perceived credibility and potential action for misinformation, with videos being particularly effective in correcting text-based misinformation. | | |
| Tsipursky et al. [184] | The Pro-Truth Pledge (PTP) has been shown to effectively reduce the sharing of misinformation and encourage truthful behavior on social media (d=-1.93). | | |

Continued from previous page

| Source | Beneficial Effects | Beneficial Perceptions | Not effective / counterproductive |
|---|---|---|---|
| Tulin et al. [185] | Truth sandwich in fact-check had indirect benefits such as more positively perceived intentions of fact-checkers and less reactance to reading subsequent fact-checks compared to classic fact check that repeats false claim (small effect sizes, e.g., for classic fact-check: $\beta$=-.13 | | Truth sandwich was not effective in correcting false beliefs but had indirect benefits. |
| Tully et al. [186] | Users tend to provide accurate information in corrections, particularly after viewing other corrections. However, users are generally unlikely to respond to tweets containing misinformation (exposure to corrections: $\eta^2$=0.001; tone of corrections: $\eta^2$=0.010). | | neither exposure to corrections nor tone of corrections increases the self-reported likelihood of responding to the misinformation tweet as compared to the misinformation-only condition |
| Tully et al. [187] | News literacy messages alter misinformation perceptions, however not with a single message (e.g., partial $\eta^2$=.0.006 for NL tweet leading participants to rate stories overall as less credible than texting tweet) | | |
| van der Meer et al. [188] | | | Warnings can prime general distrust in authentic news. |
| van der Meer and Jin [189] | Corrective information is effective in debunking misinformation, and factual elaboration compared to simple rebuttal stimulates intentions to take protective action, with government agencies and news media being more effective in improving belief accuracy compared to social peers. | | |
| Velasco et al. [190] | | The browser extension that allows to insert text and creates a (binary) feedback based on logistic regression was rated highly acceptable in terms of functionality, reliability, usability, efficiency, maintainability, and portability. | |
| Velez et al. [191] | Fact-checks undo effects of misinformation on beliefs (large and significant effect of over .26 scale points change). No Backfire effect was observed. | | |
| von der Weth et al. [192] | Nudging users toward more conscious posting and sharing behavior by using linguistic analysis to infer the factuality of content and the credibility of sources is effective in reducing the reach and speed of spread of misinformation. | | |
| Vraga et al. [196] | User corrections of a meme containing misinformation are effective in reducing the credibility assessment of the misinformation post ($\eta^2$=0.077) and misperceptions ($\eta^2$=0.088) | | Exposure to news literacy messages did not enhance the effectiveness of corrective responses or boost NL attitudes and may have generated cynicism. |
| Vraga and Bode [193] | Social corrections providing a source are effective compared to not giving a source (partial $\eta^2$=0.035) | | |
| Vraga and Bode [197] | Misinformation correction by expert group is effective without loosing the groups credibility and trustworthiness in the context of a health topic (misinformation correction: partial $\eta^2$=0.009; trustworthiness: partial $\eta^2$=0.001; credibility: partial $\eta^2$=0.004) | | misinformation corrections of a single user is not effective |
| Vraga et al. [198] | Expert organizations can be effective in successfully correcting misinformation on social media on two controversial health topics | | |
| Vraga et al. [194] | User corrections in real-time partially reduce the effect of misinformation videos on beliefs (partial $\eta^2$=0.03 compared to no intervention) but not on intentions. | | |
| Vraga et al. [199] | Logic-based and humor-based rhetorical corrections reduce misperceptions only for some topics (partial $\eta^2$=0.013). | | |
| Vraga et al. [195] | Logic-focused (before and after misinformation) and fact-focused (after the misinformation) corrections reduce misperceptions, with logic-focused corrections appearing to reduce the credibility of misinformation and fact-focused corrections being more credible. | | |
| Wahlheim et al. [200] | Reminders of misinformation are effective to diminish the negative effects of fake-news exposure short-term (d=0.29) | | |
| Waltenberger et al. [201] | Contextualizing user profiles with data from previous contributions helped users contextualize posts, identify political tendencies, distinguish humor from problematic mindsets (qualitatively measured) | | |

Continued from previous page

| Source | Beneficial Effects | Beneficial Perceptions | Not effective / counterproductive |
|---|---|---|---|
| Wang and Huang [204] | One sided narrative messages are more effective then two-sided ones for correcting misinformation on e-cigarettes | | Effect disappeared if participants had smoked e-cigarettes before |
| Wang [202] | Participants accept unwelcome fact-checks on Facebook but welcome fact-checks on Line (private messaging app). Fact-checks help increase media literacy in open platforms and hamper media literacy in private messaging apps. | | |
| Westbrook et al. [207] | External correction (news source labeling misinformation as false) influences perceptions of misinformation source. Perceptions of the misinformation source can cause changes in belief in misinformation. (a priori power analysis allowed for desired power of 0.8) | | |
| Wijnker et al. [208] | All investigated correction methods for misleading graphs were effective for debunking misinformation directly after correction and reduces over time. Showing an accurate alternative graph was more effective than visual cues or text-based warning cues to activate graph literacy or warning messages for possible deceit. | | |
| Wood et al. [210] | Debunking messages of healthcare professional lead to increase in beliefs about risks of vaccines in the UK but not the US. Messages from political authorities and discrediting messages had no effect. There is a joint importance of message source and messaging strategy regarding effectiveness of debunking (e.g., debunking by health experts reduced belief that vaccines cause severe side effects by 0.19 points on Likert scale) | | |
| Zade et al. [214] | | Tweet trajectory (e.g., unfamiliar activity invokes skepticism in following network) and contextual cues (e.g., profile description helps infer purpose of account) helped support users in assessing credibility (qualitatively evaluated). | |
| Zhao [217] | Participants exhibit a more positive attitude towards corrective messages and have higher vaccination certainty when such messages are present across multiple social media platforms, as opposed to only one platform. | | |
| Zhang et al. [216] | Concise corrections are more effective than exhaustive ones. Graphical explanation has small positive effect (e.g., Spearman's $\rho$=0.126). | | Textual explanations for why misinformation is wrong do not significantly affect effectiveness. Warnings in a tough tone make corrections worse. Textual and graphic warnings have negative associations with correction effectiveness. |
| Zheng and Ma [218] | Explanatory annotations and interactive linking in misinformation combining text and visualizations can significantly lower perceived credibility (e.g., d=-0.367). The effect to raise awareness is limited/marginal while linking was more effective than annotation (e.g., d=-0.367) | | |