TAPFLARE / RESEARCH & PRACTICAL GUIDANCE

Web Design Statistics Sources: 50 ms, 94% and 62-90% Traced

> Tapflare is a flat-rate monthly subscription for graphic design and website development. A dedicated project manager assigns each request to senior designers and developers, so companies get steady creative and web output without hiring.

Published by Tapflare · 2026-11-11 · web design statistics sources · 50 milliseconds first impression · 94% first impressions design · lindgaard 2006 · sillence 2004 · stanford web credibility · singh 2006 color · design statistics fact check · citation provenance

Original article: https://tapflare.com/articles/design-statistics-source-ledger-50-ms-94-percent



In this article

1. Executive Summary
2. Introduction and Background
3. Key Findings by Claim
4. Data Analysis and Evidence
5. Implementation Considerations and Process Changes
6. Implications and Future Directions
7. Frequently Asked Questions (FAQs)
8. Conclusion

Executive Summary

This ledger traces the design statistics most often pasted into pitch decks, briefs and landing pages back to the studies that produced them, using pages opened on 2026-11-11. The short answer for web design statistics sources is that each headline number can be traced to a study or document that states something related, though the 75% and 62-90% claims lack a traceable primary data source, and the wording in circulation is usually broader than the evidence. None of the four best-known claims can be quoted as commonly worded without qualification.

The **50 ms** claim comes from Lindgaard, Fernandes, Dudek and Brown (2006). The paper reports that visual appeal ratings of homepage screenshots shown for 50 ms agreed closely with ratings after 500 ms, and concludes that "visual appeal can be assessed within 50 ms" (Lindgaard et al. 2006). In Study 3, **40 participants** were split between the two exposure times (Lindgaard et al. 2006), and the authors state that the effect of a first impression on later behaviour such as buying was not tested (Lindgaard et al. 2006).

The **94%** figure comes from Sillence and colleagues (2004). It is the weight of design among comments explaining why sites were rapidly rejected, in a study of **fifteen women** looking for menopause information (Sillence et al. 2004). The same paper reports that **83%** of positive comments about trusted sites concerned content and **17%** design, which the authors summarize as a staged model (Sillence et al. 2004). It is not a share of "first impressions" in general.

The Stanford credibility study counted design look in **46.1%** of **2,440** coded comments, not 46.1% of people (Stanford credibility study). No primary source located in this research states "75% judge credibility by design"; the Stanford study reports design look as a share of coded comments, not of people.

The **62-90%** colour range appears in the abstract of Singh (2006), a paper whose stated method is a literature review (Singh 2006 abstract). Peer-reviewed colour reviews describe the evidence base as "nascent" (PubMed colour review).

For consumer-survey numbers the pattern repeats: Edelman's **81%** is a "deal breaker or a deciding factor" figure from a 2019 survey (Edelman survey), and Emplifi's **86%** refers to "two to three" bad experiences in a 2022 survey (Emplifi release). Section "How to Quote Safely" gives study-specific wording for each. Where a decision depends on design quality, testing the page itself is the more defensible evidence.

Introduction and Background

Design-spend arguments lean on a small set of numbers that travel from blog to blog: users judge a site in 50 milliseconds, 94% of first impressions are design related, 75% judge credibility by how a site looks, and colour drives 62-90% of snap judgments. A marketing lead, founder or account manager who copies one of these into a deck inherits whatever distortion accumulated on the way. This report is a provenance ledger: for each claim it identifies the original study, what was measured, on whom, and which wording the evidence supports. It is written for the person deciding whether a number can be quoted without embarrassment, and how to phrase it. It does not offer new survey data, benchmarks or case-study figures, and it does not argue that design is unimportant.

How to read this ledger

Three terms are used consistently.

- **Primary source:** the paper, report or sponsor publication in which the authors of a study describe their own data. A restatement in a review, press release distribution or blog is secondary.
- **Secondary source:** any later page that reports the number. Secondary pages are used here only to show what wording circulates or to locate an origin.
- **Unsupported:** no primary source located in this research states the claim as worded. This is a finding about this search, not proof that no such data exist.

Each claim receives one of three verdicts: **supported as worded**, **supported with narrower wording**, or **not supported as worded**. Retrieval date for every page is 2026-11-11. Several primary documents could not be read in full, and the limits are stated where they matter: the Emerald full text of Singh (2006) is paywalled and only its abstract and deposited reference list were read (emerald.com); Edelman's full 2019 report sits behind a registration form (info.edelman.co.uk); and Emplifi's report is gated (go.emplifi.io). This ledger is a provenance reference rather than a roundup of design statistics.

Why provenance matters for a buying decision

A statistic's strength depends on its design. A rating of screenshots in a laboratory, a weighting of focus-group comments, a vendor-commissioned poll and a literature-review sentence are different kinds of evidence. Library guidance on evaluating statistics asks the same questions each time: whether the original source can be found (guides.smu.edu), what the sample size and sampling method were (libguides.unbc.ca), and what context is missing in re-reporting (libguides.sbcc.edu). The ledger below applies those questions to the design numbers.

Key Findings by Claim

Table 1 below summarizes the ledger: the wording in circulation, the origin, what was actually measured, the verdict and a safer sentence. Details and sources for each row follow in the subsections.

Wording in circulation	Original study	Sample and method	What it measured	Verdict	Safer wording
Users form an opinion of a website in 50 ms	Lindgaard et al. (2006), Behaviour & Information Technology 25(2):115-126 (typeset.io)	Study 1: 100 homepage screenshots; Study 3: 40 participants, 50 ms vs 500 ms (www.makinggood.ac.nz)	Visual appeal ratings of screenshots; agreement between 50 ms and 500 ms ratings	**Supported with narrower wording**	"In a 2006 lab study, appeal ratings of homepage screenshots shown for 50 ms agreed closely with ratings after 500 ms."
94% of first impressions are design related	Sillence et al. (2004), CHI 2004 (dl.acm.org)	Fifteen women, menopause information, focus groups (hci.rwth-aachen.de)	Weighting of comments about why sites were rejected or mistrusted (hci.rwth-aachen.de)	**Supported with narrower wording**	"In a 15-person study of health sites, 94% of weighted comments about rejecting a site concerned design."
75% judge credibility by design	No primary statement located	No matching figure located in the Stanford sources examined	Not measured as worded	**Not supported as worded**	Avoid, or cite the 46.1% comment share.
Design look drives 46.1% of credibility judgments	Fogg et al. (2003) and 2002 Stanford/Consumer WebWatch report (dl.acm.org)	2,684 people; 2,440 coded comments; non-representative (credibility.stanford.edu)	Share of comments mentioning design look	**Supported with narrower wording**	"46.1% of coded comments mentioned design look."
Colour accounts for 62-90% of snap judgment	Singh (2006), Management Decision 44(6):783-789 (emerald.com)	Literature review; no original data (emerald.com)	A range stated in the abstract; the full text was not read	**Not supported as worded**	"A 2006 literature review states that colour is a large part of quick product assessments."
81% of consumers trust brands before buying	Edelman special report (2019) (edelman.com)	16,000 people in eight countries (edelman.com)	Trust as a major consideration, ranked behind quality, convenience, value and ingredients	**Supported with narrower wording**	"In a 2019 Edelman survey, 81% said brand trust is a deal breaker or deciding factor."
86% leave after two bad experiences	Emplifi, with Harris Interactive (2022) (www.businesswire.com)	More than 2,000 US and UK consumers (www.businesswire.com)	Intent to leave after "two to three" bad service experiences	**Supported with narrower wording**	"In a 2022 vendor-commissioned US and UK survey, 86% said they would leave after two to three bad service experiences."

Wording in circulation	Original study	Sample and method	What it measured	Verdict	Safer wording
85% say colour is the primary reason to buy; 93% say visual appearance decides	Closest located: Seoul International Color Expo 2004, reported on a colour consultancy's page	Method and sample size not located	84.7% said colour is "more than half" of purchase factors (www.colorcom.com); 92.6% put most importance on visual factors	**Not supported as worded**	Avoid, or attribute to the 2004 Expo survey with the original phrasing.

The pattern across the table is consistent. Five of the eight rows have a real origin whose scope is narrower than the sentence in circulation, and three have no traceable primary statement of the number as it is quoted. Nothing in the table shows that design has no effect; the point is that each study answers a smaller question than the one the circulating sentence implies.

"50 milliseconds": Lindgaard et al. (2006)

What the paper did. The abstract states that three studies were run to ascertain how quickly people form an opinion about web page visual appeal (typeset.io). Study 1 used screenshots of 100 web pages (www.makinggood.ac.nz), collected as best and worst examples of visual appeal (www.makinggood.ac.nz). Study 2 narrowed the set to the 25 highest-rated and 25 lowest-rated pages (www.makinggood.ac.nz). Study 3 repeated the 500 ms condition and added a 50 ms condition (www.makinggood.ac.nz), with a sample of 40 participants (www.makinggood.ac.nz). Ratings at 50 ms correlated closely with ratings at 500 ms: the paper reports r = .947 for the first phase and r = .953 for the second phase of the 50 ms versus 500 ms comparison, and describes the first-phase result as slightly lower than, but comparable to, the Study 2 correlation for the 500 ms condition alone (www.makinggood.ac.nz). The conclusion is that appeal ratings were as robust at 50 ms as at 500 ms (www.makinggood.ac.nz).

What the paper did not claim. Four limits are stated in the paper or in replications.

- **Not a threshold.** The authors say their aim is not to determine an accurate threshold (www.makinggood.ac.nz).
- **Extreme stimuli.** Pages between the appealing and unappealing extremes in Study 1 were eliminated in Studies 2 and 3 (www.makinggood.ac.nz).
- **Lower agreement at 50 ms.** The authors note that agreement was lower at 50 ms and speculate that at 500 ms participants took in more content (www.makinggood.ac.nz).
- **No behaviour measured.** The paper's conclusion says research is needed on the effect of a first impression on later behaviour, such as selecting a site or buying (www.makinggood.ac.nz), and that more research is needed to address that question (www.makinggood.ac.nz).

Replications and refinements. Tuch and colleagues rated 119 website screenshots at presentation times of 50, 500 and 1000 ms (research.google.com), with 59 participants in the first study (research.google.com) and a second study of 82 participants (research.google.com) that shortened the times to 17, 33 and 50 ms. They report that visual complexity and prototypicality affect aesthetic ratings even within 17 ms (research.google.com), and they used a masking stimulus to control exposure (research.google.com). Thielsch

and Hirschfeld (92 participants) (uni-muenster.de) note that the early studies lacked masking and used repeated stimulus presentation (uni-muenster.de), and report that a later study found a lower correlation of r = .73 (uni-muenster.de). A third paper they discuss suggests that early impressions' influence on later judgments could be less stable than previously believed (uni-muenster.de). Tractinsky and colleagues ran their own replication, in which 40 participants evaluated 50 web pages in two phases (www.sciencedirect.com), and noted that intra-user consistency was considerably lower in the 50 ms condition of the original than in the 0.5 s condition (www.sciencedirect.com).

A note on restated numbers. Secondary summaries restate the Study 3 correlation slightly differently: Thielsch and Hirschfeld describe correlations between both conditions as remarkably high (uni-muenster.de), while Tuch et al. summarize the last experiment as a "very high" correlation (research.google.com). The originating paper is the place to quote from.

Limits on real-world generalization. Tuch et al. caution that passive screenshot viewing cannot be associated directly with real-life perception (research.google.com), that real interactions matter (research.google.com), and that it remains to be seen whether results hold for other kinds of sites (research.google.com). News coverage in 2006 summarized the work as impressions being made in the first 50 milliseconds of viewing (www.nature.com), and the idea has since appeared in a Google Research blog as "In less than 50 milliseconds, users build an initial" impression (research.google).

Verdict: supported with narrower wording. It is a finding about appeal ratings of screenshots in a lab, not about purchase intent, bounce rate or conversion. Behavioural data on real sessions point to a longer window: Nielsen Norman Group reports that the average page visit lasts a little less than a minute (www.nngroup.com) and that the first 10 seconds are critical (www.nngroup.com).

"94% of first impressions are design related": Sillence et al. (2004)

What the paper did. The CHI 2004 paper observed fifteen women faced with a risky health decision (hci.rwth-aachen.de). The Method section describes fifteen women at various stages of the menopause (hci.rwth-aachen.de), aged 41 to 60 with a mean of 49 (hci.rwth-aachen.de). Each participant attended four 2-hour sessions (hci.rwth-aachen.de), and discussions and verbal protocols were transcribed and content-analysed (hci.rwth-aachen.de). A peer-reviewed follow-up in Social Science & Medicine confirms the same fifteen women with menopause and hormone replacement therapy decisions (researchportal.northumbria.ac.uk).

Where the 94% sits. The paper's passage on rejection begins with participants' first impressions of a website (hci.rwth-aachen.de). Table 2 is captioned as factors relating to the rejection and mistrust of websites (hci.rwth-aachen.de), and its final column shows the number of times a factor was mentioned (hci.rwth-aachen.de). The authors write that, as Table 2 shows, the main reason websites were rejected was design (hci.rwth-aachen.de). The design rows total 94% and the lone content row, irrelevant or inappropriate content, is 6% (hci.rwth-aachen.de). The authors' later article reprints the split as design factors at 94% of cited reasons (www.researchgate.net). The base is comments about rejection; the paper does not give raw counts in the text read for this report, so the weighting cannot be converted to a headcount.

The opposite split for trusted sites. Table 3, captioned as positive features about selected, trusted sites (hci.rwth-aachen.de), shows content at 83% and design at 17% (www.researchgate.net). The authors state that content factors were more important there (hci.rwth-aachen.de).

The authors' own conclusion. The paper proposes a staged model of trust (hci.rwth-aachen.de), in which visual appeal influences early decisions to reject or mistrust sites (hci.rwth-aachen.de), with content credibility mattering at the selection stage. A later framework paper summarizes the same pattern across studies: design appeal predicted rejection (eprints.whiterose.ac.uk), and the Social Science & Medicine abstract describes women sometimes rejecting high-quality content because of poor design (researchportal.northumbria.ac.uk).

Different sample, different number. A 2005 poster on a separate sample of thirteen hypertension patients (archives.iw3c2.org) reports that 83% of rejection comments concerned design (archives.iw3c2.org) and that 74% of comments about selected or trusted sites concerned content (archives.iw3c2.org). The 94% is therefore specific to the menopause study. The authors flag one caveat: whether participants in a focus group might be tempted to give socially desirable responses (hci.rwth-aachen.de).

Verdict: supported with narrower wording. Safe: "In a 2004 study of fifteen women searching for menopause information, 94% of weighted comments about why they rejected sites concerned design." Not safe: any sentence that calls 94% a share of first impressions, users, or sites in general, or that omits the 83% content finding for trusted sites.

"75% judge credibility by design" and "46.1%": the Stanford work

The study behind 46.1%. Fogg and colleagues' DUX 2003 paper reports that 2,684 people evaluated the credibility of two live Web sites on a similar topic (dl.acm.org), with design look present in 46.1% of the comments (dl.acm.org). The Stanford publications page describes the paper as involving "over 2,500 participants" (credibility.stanford.edu), which is consistent with the 2,684 figure. The full report, released 29 October 2002 (credibility.stanford.edu), collected data from 15 June to 15 August 2002 (credibility.stanford.edu), and states that a total of 2,684 people completed the study (credibility.stanford.edu). It reports design look in 46.1% of the 2,440 comments (credibility.stanford.edu). Two independent coders assigned codes, and a comment could receive more than one code (credibility.stanford.edu), so category percentages overlap and are shares of comments, not of people. The authors say their charity recruiting method does not provide a representative sample of Web users (credibility.stanford.edu).

Category variation. Design-look comments were more frequent for finance (54.6%) and search engines (52.6%) than for health (41.8%) (credibility.stanford.edu). The second-largest category was information design and structure at 28.5% (credibility.stanford.edu).

Where "75%" could come from. A search of the DUX abstract, the 2002 report, the Stanford guidelines page and the earlier Stanford-Makovsky study found no sentence stating that 75% of users judge credibility by design. The earlier study had 1481 participants (credibility.stanford.edu) and a different method, and the CHI 2001 study had over 1400 people evaluating 51 elements (dl.acm.org).

The guidelines wording. Stanford's guidelines say that people quickly evaluate a site by visual design alone (credibility.stanford.edu), based on three years of research that included over 4,500 people (credibility.stanford.edu). This is a design guideline, not a percentage.

A separate experiment. Robins and Holmes had 20 subjects (makinggood.ac.nz) make 840 credibility judgments (makinggood.ac.nz) and found the higher-aesthetic version of the same content rated more credible in 19 of 21 cases (makinggood.ac.nz). That is an experimental result with a small, mostly graduate-student

sample, and it is a safer citation than the 75% for the claim that appearance affects credibility.

Verdict. 46.1% of comments: supported with narrower wording. 75% of users: not supported as worded.

"Colour accounts for 62-90% of snap judgment": Singh (2006)

What the abstract says. The Emerald abstract states that people make up their minds within 90 seconds of initial interactions with people or products (emerald.com), and that "about 62-90 percent of the assessment is based on colors alone" (emerald.com). The paper describes its method as a review of the literature on colour psychology in a marketing context (emerald.com), and the abstract itself highlights inconsistencies and controversies in that literature (emerald.com).

Tracing the range. Only the abstract and deposited reference list of the paper were read for this report. Because the full text is paywalled, what the body text cites for the range could not be checked; the reference list alone cannot say which entry, if any, supports it.

The same range on another page. A colour consultancy's page states that research reveals people make a subconscious judgment within 90 seconds of initial viewing (www.colorcom.com) and that between 62% and 90% of that assessment is based on colour alone (www.colorcom.com). Its attribution line names the Institute for Color Research (www.colorcom.com).

What peer-reviewed reviews say. Elliot and Maier's 2014 review states that the literature remains at a nascent stage (pubmed.ncbi.nlm.nih.gov) and that considerable work on boundary conditions, moderators and real-world generalizability is needed (pubmed.ncbi.nlm.nih.gov). Elliot's 2015 follow-up recommends patience and prudence regarding conclusions about real-world application (pmc.ncbi.nlm.nih.gov). Labrecque, Patrick and Milne note relatively little research in marketing (api.crossref.org), and Labrecque and Milne state that little academic research has investigated the role colour plays in marketing (link.springer.com); their own paper reports four studies on brand perceptions rather than a share of snap judgments (link.springer.com).

The 85% and 93% variants. The colour consultancy's page reproduces a 2004 Seoul International Color Expo survey, in which 92.6% said they put most importance on visual factors when purchasing and 84.7% said colour accounts for more than half (www.colorcom.com). Those two numbers are the nearest located to the circulating "93% visual appearance" and "85% colour" claims, and rounding is consistent with them, but no source links the two explicitly, so the connection is a hypothesis. The Expo's method and sample size were not found, and the original was not reachable.

Verdict: not supported as worded. "Colour accounts for 62-90% of snap judgments" is not a measured result reported in the abstract, which describes a literature review. A defensible version names Singh (2006) as a review and states that colour influences perception, which the reviews support, without a percentage.

Data Analysis and Evidence

The consumer-survey statistics follow a different pattern from the academic ones: the originals are vendor or consultancy surveys, often gated, and the circulating wording compresses a more specific question. Each entry below states the sponsor, year, population and question where they could be read.

Consumer-survey figures traced to their sponsors

- **81% brand trust (Edelman, 2019).** Edelman's press release says brand trust (81 percent) is a deal breaker or a deciding factor (edelman.com), behind quality (85%), convenience (84%), value (84%) and ingredients (82%) (edelman.com), in a survey of 16,000 people in eight countries (edelman.com). The full report requires registration (info.edelman.co.uk), so the questionnaire could not be inspected.
- **86% leave after bad experiences (Emplifi, 2022).** The release says 86 percent will leave a brand they were once loyal to after only two to three bad customer service experiences (www.businesswire.com), surveying more than 2,000 consumers in the US and UK (www.businesswire.com). The report page says the research was by Harris Interactive and commissioned by Emplifi (go.emplifi.io) and words the finding as over 4 in 5 respondents leaving after three or fewer instances of poor service (go.emplifi.io). The figure is therefore tied to one survey's question wording and year.
- **Other "bad experience" figures.** PwC's 2017/18 survey of 15,000 people (www.pwc.com) reports that 59% in the US will walk away after several bad experiences (www.pwc.com), and that a share will walk away after just one (www.pwc.com). A different 86% appears in a 2009 RightNow and Harris report: 86% of consumers quit doing business with a company after a bad customer experience (www.slideshare.net), and a 2012 Oracle brief reports 86 percent in a pay-more context (www.oracle.com). Several distinct "86%" statistics exist, so the sponsor and year must accompany the number.
- **Brand consistency (Lucidpress, 2016 and 2019).** The 2016 release describes 23 percent as the average revenue increase estimate attributed to always presenting the brand consistently (www.prweb.com), among organizations rather than consumers. The 2019 release says Lucidpress surveyed over 200 organizations (prnewswire.com), and the report states that constituents reported an average of 33% growth if they maintained consistency (pubsecure.lucidpress.com). These are self-estimates by organizations. The report's 43% concerns about 43% of the survey pool reporting inconsistent branding (pubsecure.lucidpress.com), not 43% of consumers distrusting inconsistent brands. A 2026 vendor survey of 200 consumers in the US and Europe reports that nearly three quarters are less likely to buy from a brand that feels inconsistent (www.storyblok.com).
- **38% unattractive layout (Adobe, 2015).** The Adobe State of Content report, read through a mirror of the original PDF because the Adobe URL was unreachable, lists unattractive layout or imagery with 35% switching devices and 38% stopping engaging (delovoymir.biz). It surveyed 2,008 consumers (delovoymir.biz), was commissioned by Adobe and produced by Edelman Berland (appdevelopermagazine.com), and asked what would make people switch devices or stop viewing content on any device (delovoymir.biz). It is stated behaviour about content, not observed behaviour about websites.
- **88% less likely to return (Gomez, 2009 fieldwork).** MarketingProfs reports that 88% are less likely to return to the poor-performing site (www.marketingprofs.com), among people who had purchased online in the previous nine months (www.marketingprofs.com), with interviews conducted December 16-22, 2009 (www.marketingprofs.com). A second outlet reports the same figure (www.pcworld.com). The question concerned site performance during peak traffic, not visual design.

Table 2 below reconciles studies that carry more than one number for the same finding, which is the most common source of confusion when two writers quote the same study differently.

Study	Number A	Number B	Why both exist
Sillence et al. (2004)	94% design among rejection comments (www.researchgate.net)	83% content among trusted-site comments (www.researchgate.net)	Two different tables answering two different questions in one paper (hci.rwth-aachen.de)
Sillence follow-ups	94% design in menopause sample (www.researchgate.net)	83% design in rejection comments in a hypertension sample (archives.iw3c2.org)	Different samples; 15 women vs 13 patients (archives.iw3c2.org)
Stanford DUX 2003	46.1% of comments mention design look (credibility.stanford.edu)	28.5% mention information design (credibility.stanford.edu)	Stanford counts shares of coded comments (codes can overlap); the figures are not shares of people
Stanford participants	"over 2,500" in the title (credibility.stanford.edu)	2,684 in abstract and report (dl.acm.org)	Publications-page description versus completed sample in the report
Lindgaard et al. (2006)	r about .95 between 50 ms and 500 ms (www.makinggood.ac.nz)	r = .73 in a later study as summarized by Thielsch and Hirschfeld (uni-muenster.de)	Different stimuli and procedures; later work used masking (uni-muenster.de)
Singh (2006)	62-90 percent of assessment (emerald.com)	Abstract and reference metadata read; full text not read	A review abstract versus a measured result
Emplifi (2022)	86% after "two to three" bad experiences (www.businesswire.com)	"three or fewer" in the report page (go.emplifi.io)	Report page uses different wording from the press release

The reconciliation shows that disagreement between numbers is usually explained by the question asked, the base of the percentage and the sample, not by error. For a writer, the useful habit is to quote the number together with its base: comments, respondents, organizations or pages.

What the evidence supports overall

Taken together, the sources support a modest and defensible set of statements. Appeal judgments of screenshots form very quickly in lab conditions (www.makinggood.ac.nz). In some contexts, design features figure prominently in comments about why sites are rejected (hci.rwth-aachen.de) and in comments about credibility (credibility.stanford.edu). Appearance and perceived usability are linked: Nielsen Norman Group describes the aesthetic-usability effect as the tendency to perceive attractive products as more usable (www.nngroup.com), citing a 1995 study of 26 ATM interface variations with 252 participants (www.nngroup.com). A 2000 study with 132 third-year engineering students found that actual usability had no effect on post-use perceptions, while aesthetics did affect perceived usability (www.ise.bgu.ac.il). All of these are narrower than "design drives X% of anything", and the largest consumer surveys measure stated intent, not behaviour.

Implementation Considerations and Process Changes

How to quote safely

Three sentence patterns cover most needs. *Table 3* below shows each pattern with a worked example per claim.

Pattern	Template	Example built from this ledger
Study-specific	"In [year], [authors] found [result] among [sample] ([venue])."	"In 2006, Lindgaard and colleagues found that appeal ratings of homepage screenshots shown for 50 ms correlated closely with ratings after 500 ms (Behaviour & Information Technology)." (www.makinggood.ac.nz)
Hedged	"Research suggests [effect], although [limitation]."	"Research on first impressions suggests appeal judgments form within a fraction of a second, although the studies used screenshots and did not measure purchases." (www.makinggood.ac.nz)
Avoid	A bare percentage with no base, sample or year.	Avoid: "94% of first impressions are design related." Use instead: "In a 2004 study of 15 women, 94% of weighted comments about rejecting health sites concerned design, while 83% of comments about trusted sites concerned content." (www.researchgate.net)

The templates work because each carries the three details a reader needs to judge the number: who was studied, what was counted and when. They also keep the sentence true if the reader opens the source.

A tracing routine for any design statistic

Before a number enters a deck or brief, the steps below take about fifteen minutes per number.

- **Find the first appearance.** Follow links upstream until reaching a paper, report or sponsor page. Library guidance recommends identifying whether a source is original or re-reporting (libguides.sbcc.edu).
- **Record the base.** Note whether the percentage is of people, comments, pages or organizations, and the sample size (libguides.unbc.ca).
- **Check the question.** Stated intent ("would you leave") and observed behaviour are different kinds of evidence. Nielsen Norman Group advises attending to what users do as well as what they say (www.nngroup.com).
- **Check the date and population.** The 2004 menopause study (hci.rwth-aachen.de) and the 2002 charity-recruited sample (credibility.stanford.edu) are old and specific.
- **Note the gating.** If the report is gated, say "per the sponsor's press release" rather than implying the questionnaire was reviewed.
- **Choose the verdict.** Supported as worded, supported with narrower wording, or not supported as worded, and write the sentence to match.

Process changes for teams that publish statistics

- **Keep a source column.** Every statistic in a deck gets a row with the primary URL, retrieval date and wording approved.

- **Prefer ranges of evidence to single numbers.** Two independent studies with different methods carry more weight than one figure repeated.
- **Separate stated from observed.** Mark survey intent figures as such.
- **Re-check annually.** Gated reports move, and sponsors may publish new surveys with different wording.

Implications and Future Directions

When a statistic should not carry the decision

Whether design work is worth commissioning is a question about a specific site, audience and goal. None of the studies above measured revenue, conversion or retention for a business like the reader's. A statistic can motivate a hypothesis, but it cannot substitute for a measurement. Several measurable alternatives exist.

- **Test your own pages.** Nielsen Norman Group's guidance is that the best results come from testing no more than 5 users and running as many small tests as affordable (www.nngroup.com). This answers whether people in a given audience understand and like a page.
- **Use A/B tests for outcomes, qualitative tests for reasons.** Nielsen Norman Group notes that in an A/B test, all you know is that, statistically, more people performed a certain action (www.nngroup.com); why requires other methods.
- **Measure performance directly.** Google's web.dev defines good-experience thresholds for loading, interactivity and visual stability, for example that Largest Contentful Paint should occur within a stated time of the page starting to load (web.dev). These are measurable page-experience metrics, unlike a percentage of "judgment".
- **Measure colour accessibility.** WCAG 2.2 success criterion 1.4.3 requires a contrast ratio of at least 4.5:1 for text, with listed exceptions (www.w3.org), and states that computed values should not be rounded (www.w3.org). A contrast ratio can be checked on any page; a "62-90%" colour share cannot.
- **Look at real dwell times.** Microsoft researchers modelled page abandonment as a failure process in reliability analysis (www.microsoft.com), which gives a way to think about how long people stay that is closer to behaviour than a 50 ms screenshot rating.

For teams that need design capacity rather than more evidence, the practical question becomes how to get design and web work done, for example through in-house staff, freelancers, agencies or subscription providers. Nothing in this ledger depends on that choice.

Future directions

- **Behavioural replication.** The Lindgaard authors identify the effect of first impressions on later behaviour as an open question (www.makinggood.ac.nz). Field experiments that connect appeal to measured outcomes on real sites would fill the largest gap.
- **Non-health samples.** Sillence's staged model comes from health information seeking (hci.rwth-aachen.de), and Tuch et al. say it remains to be seen whether results transfer to other site types (research.google.com). Replications in e-commerce and SaaS would test the generalization.
- **Colour evidence.** Reviews call the colour literature nascent (pubmed.ncbi.nlm.nih.gov) and ask for work on boundary conditions and real-world generalizability (pubmed.ncbi.nlm.nih.gov).

- **Open questionnaires.** Gated vendor reports limit verification; publishing questions and bases would make survey statistics easier to quote correctly.

Frequently Asked Questions (FAQs)

Who said users form an opinion of a website in 50 milliseconds?

Gitte Lindgaard, Gary Fernandes, Cathy Dudek and J. Brown, in a 2006 paper in Behaviour & Information Technology titled "Attention web designers: You have 50 milliseconds to make a good first impression!" (typeset.io). The 50 ms source for the website first-impression claim is therefore a study of appeal ratings of screenshots, with the conclusion that visual appeal can be assessed within 50 ms (www.makinggood.ac.nz).

What did Lindgaard 2006 actually test?

It tested ratings of homepage screenshots: 100 pages in Study 1 (www.makinggood.ac.nz), and 50 extreme-appeal pages in Studies 2 and 3 with 40 participants split across 50 ms and 500 ms in Study 3 (www.makinggood.ac.nz). It did not test purchase intent, abandonment or conversion (www.makinggood.ac.nz).

What is the source of "94% of first impressions are design related"?

Sillence, Briggs, Fishwick and Harris (2004), CHI 2004, "Trust and Mistrust of Online Health Sites" (dl.acm.org). The 94% is the weight of design among comments about why menopause-information sites were rejected, from fifteen women (hci.rwth-aachen.de). The same paper reports 83% content for trusted sites (www.researchgate.net).

Is it true that 75% of users judge credibility by design?

No primary source located states this. The Stanford comment study reports design look in 46.1% of 2,440 coded comments (credibility.stanford.edu).

Where does "colour accounts for 62-90% of snap judgment" come from?

It appears in the abstract of Singh (2006), Management Decision 44(6):783-789 (emerald.com), a literature review (emerald.com). The same range is also stated on a colour consultancy's page.

Does the Singh 2006 "90 percent" figure come from an experiment?

No. The abstract's method is a review of the literature (emerald.com), and it states that people make up their minds within 90 seconds (emerald.com). The percentage is not reported as a measured result of the paper itself.

Is there a fact check of these website first-impression statistics?

This ledger is one. The verdicts are in *Table 1*: the 50 ms and 94% claims are supported with narrower wording, and the 75% and 62-90% claims are not supported as worded.

Which design number is safest to quote?

The most defensible are the study-specific ones with their base: the Robins and Holmes experiment (19 of 21 pairs rated more credible with the higher-aesthetic version) (makinggood.ac.nz) and the Lindgaard correlation described with its screenshot context (www.makinggood.ac.nz).

Conclusion

Each of the four headline design statistics can be traced to a study or document that states something related, though the 75% and 62-90% claims lack a traceable primary data source, and none survives the usual retelling intact. The **50 ms** result concerns appeal ratings of screenshots in a lab. The **94%** is a comment weighting from fifteen women reading menopause sites, paired in the same paper with an 83% content finding for trusted sites. The **46.1%** is a share of coded comments from a non-representative sample, and the **75%** has no primary statement. The **62-90%** colour range is a sentence in the abstract of a literature review. Consumer-survey numbers are vendor or consultancy polls whose wording shifts between the questionnaire, the headline and the blog post.

For a marketing lead, founder or account manager, the practical rule is to quote the study, not the percentage: who, how many, what was counted, and when. Where a decision about design spend is at stake, measured results from the reader's own pages, such as small usability tests, A/B outcomes, page-experience metrics and contrast checks, are stronger evidence than any number in this ledger.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://www.makinggood.ac.nz/media/1265/lindegaardetal_2006_attention.pdf
2. https://hci.rwth-aachen.de/materials/conferences/CHI2004/1p663.pdf
3. http://credibility.stanford.edu/pdf/How_Do_People_Evaluate_a_Web_Site's_Credibility_v37.pdf
4. https://emerald.com/insight/content/doi/10.1108/00251740610673332
5. https://pubmed.ncbi.nlm.nih.gov/23808916
6. https://edelman.com/news-awards/only-one-third-of-consumers-trust-most-of-the-brands-they-buy
7. https://www.businesswire.com/news/home/20220202005525/en/86-Percent-of-Consumers-Will-Leave-a-Brand-They-Trusted-After-Only-Two-Poor-Customer-Experiences
8. https://info.edelman.co.uk/in-brands-we-trust
9. https://go.emplifi.io/us-uk-customer-expectations-report.html
10. https://guides.smu.edu/statistics
11. https://libguides.unbc.ca/data/evaluate
12. https://libguides.sbcc.edu/statistics/evaluate
13. https://typeset.io/papers/attention-web-designers-you-have-50-milliseconds-to-make-a-3axej7ku89
14. https://dl.acm.org/doi/10.1145/985692.985776
15. https://dl.acm.org/doi/10.1145/997078.997097

16. https://www.colorcom.com/why-color-matters
17. https://research.google.com/pubs/archive/38315.pdf
18. https://uni-muenster.de/OWMS/uploads/drafts_thielsch/pdf/thielsch_hirschfeld_2012.pdf
19. https://www.sciencedirect.com/science/article/abs/pii/S1071581906000863
20. https://www.nature.com/news/2006/060109/full/news060109-13.html
21. https://research.google/blog/users-love-simple-and-familiar-designs-why-websites-need-to-make-a-great-first-impression
22. https://www.nngroup.com/articles/how-long-do-users-stay-on-web-pages
23. https://researchportal.northumbria.ac.uk/en/publications/how-do-patients-evaluate-and-make-use-of-online-health-informatio
24. https://www.researchgate.net/publication/6477584_How_Do_Patients_Evaluate_and_Make_Use_of_Online_Health_Information
25. https://eprints.whiterose.ac.uk/9905
26. https://archives.iw3c2.org/www2005/proceedings/docs/p1026.pdf
27. https://credibility.stanford.edu/publications.html
28. https://credibility.stanford.edu/pdf/Stanford-MakovskyWebCredStudy2002-prelim.pdf
29. https://dl.acm.org/doi/10.1145/365024.365037
30. https://credibility.stanford.edu/guidelines/index.html
31. https://makinggood.ac.nz/media/1276/robinsholmes_2008_aestheticsandcredibilityinwebdesign.pdf
32. https://pmc.ncbi.nlm.nih.gov/articles/PMC4383146
33. https://api.crossref.org/works/10.1002/mar.20597
34. https://link.springer.com/article/10.1007/s11747-010-0245-y
35. https://www.pwc.com/us/en/advisory-services/publications/consumer-intelligence-series/pwc-consumer-intelligence-series-customer-experience.pdf
36. https://www.slideshare.net/RightNow/right-nows-4thannualcustomerexperienceimpactreport
37. https://www.oracle.com/assets/cx-service-exec-strategy-brief-1856861.pdf
38. https://www.prweb.com/releases/2016/11/prweb13829874.htm
39. https://prnewswire.com/news-releases/study-finds-companies-with-consistent-branding-can-see-up-to-33-increase-in-revenue-300967219.html
40. https://pubsecure.lucidpress.com/6e525791-56c1-4b3b-b9c4-e3db88727507/document.pdf
41. https://www.storyblok.com/mp/inconsistent-brands
42. https://delovoymir.biz/res/upload/columns/state-of-content-oct.pdf
43. https://appdevelopermagazine.com/adobe-report-shows-growing-rise-in-consumer-distrust-of-mobile-content
44. https://www.marketingprofs.com/charts/2010/3399/poor-web-performance-hurts-brands
45. https://www.pcworld.com/article/516685/article-4802.html
46. https://www.nngroup.com/articles/aesthetic-usability-effect
47. http://www.ise.bgu.ac.il/faculty/noam/papers/00_nt_ask_di_iwc.pdf
48. https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users
49. https://www.nngroup.com/articles/putting-ab-testing-in-its-place
50. https://web.dev/articles/vitals

51. https://www.w3.org/WAI/WCAG22/Understanding/contrast-minimum.html
52. https://www.microsoft.com/en-us/research/publication/understanding-web-browsing-behaviors-through-weibull-analysis-of-dwell-time

THE PUBLISHER

About Tapflare

Tapflare gives companies an on-demand design and web team for a flat monthly fee.

How it works

Each subscriber has a dedicated project manager who learns the brand, assigns every request to a senior specialist in that kind of work, and reviews the result before it is delivered. Requests are submitted and discussed in the Tapflare portal, each plan includes a daily allotment of design or development time, and the client keeps full ownership of the export and source files.

What it covers

Graphic design, custom illustration, presentation design, logos and brand guides, video editing and motion graphics, website and mobile-app mockups, and website development, from design mockups to HTML, to drag-and-drop builders or custom code.

What it is not

Tapflare is not a marketing agency, a media-buying, SEO or copywriting service, a software product or a freelance marketplace. Where a reader needs one of those, the honest answer is to say so.

Visit

See Tapflare and its pricing.

Articles published here explain general practice and cite their sources. Prices and plan contents, including Tapflare's own, change; check the provider's current pricing page before deciding.

Website: https://tapflare.com

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.