

UNPUBLISHEDUNITED STATES COURT OF APPEALS
FOR THE FOURTH CIRCUIT

No. 19-1674

COMBE INCORPORATED,

Plaintiff – Appellee,

v.

DR. AUGUST WOLFF GMBH & CO. KG ARZNEIMITTEL,

Defendant – Appellant.

Appeal from the United States District Court for the Eastern District of Virginia, at Alexandria. T. S. Ellis, III, District Judge. (1:17-cv-00935-TSE-MSN)

Argued: January 28, 2021

Decided: April 13, 2021

Before AGEE, THACKER and QUATTLEBAUM, Circuit Judges.

Affirmed by unpublished per curiam opinion.

ARGUED: Ross Quinn Panko, ARENT FOX LLP, Washington, D.C., for Appellant. Douglas Anthony Rettew, FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP, Washington, D.C., for Appellee. **ON BRIEF:** Michael A. Grow, Laura E. Zell, ARENT FOX LLP, Washington, D.C., for Appellant. Anna B. Naydonov, Sydney N. English, FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP, Washington, D.C., for Appellee.

Unpublished opinions are not binding precedent in this circuit.

PER CURIAM:

This case involves a trademark registration dispute between two makers of vaginal-health and hygiene products that are marketed under the respective brands “Vagisil” and “Vagisan.” Combe, Inc. (“Combe”), which owns the Vagisil mark, contended that registration of Dr. August Wolff GMBH & Co. KG Arzneimittel’s (“Wolff”) Vagisan mark would likely confuse consumers, and thus should be disallowed under the Lanham Act, 15 U.S.C. § 1051 *et seq.* The district court agreed with Combe. Wolff appeals. For the reasons set out below, we affirm the judgment of the district court.

I.

Since the mid-1970s Combe has sold an array of women’s vaginal-health products under the name “Vagisil.” In 1978, it registered “Vagisil” with the U.S. Patent and Trademark Office (“PTO”) and has maintained ownership of a Vagisil mark covering “pharmaceutical preparations,” particularly “medicated cremes,” ever since. J.A. 1336. In subsequent years, Combe acquired ownership of additional registered Vagisil marks to cover its expanded product line, which now includes vaginal-health powders, wipes, washes, and moisturizers.

Since 1998 Wolff, a German-based company, has used the name “Vagisan” to sell vaginal-health and hygiene products internationally. It owns several foreign trademark registrations for Vagisan, but has not sold Vagisan products in the United States. Desiring to enter the previously untapped American market, Wolff applied with the PTO to register

the unadorned, standard characters “VAGISAN” as a U.S. trademark for various “pharmaceutical preparations” and other related products in 2012. J.A. 268.

Combe opposed the registration of Wolff’s mark, asserting Vagisan’s similarity to Vagisil would create a likelihood of confusing consumers. The Trademark Trial & Appeals Board (“TTAB”) dismissed Combe’s opposition and allowed Wolff to register the Vagisan mark.

Thereafter, Combe filed a timely civil action under 15 U.S.C. § 1071(b) seeking to have Wolff’s Vagisan mark cancelled.¹ After a bench trial, the district court found that because the Vagisan mark presented a likelihood of confusing consumers, its registration violated the Lanham Act. The district court entered judgment in favor of Combe, reversed the TTAB ruling dismissing Combe’s opposition, and ordered refusal of Wolff’s application to register the Vagisan mark.

Wolff noted a timely appeal, and the Court has jurisdiction under 28 U.S.C. § 1291 and 15 U.S.C. § 1121.

¹ This provision of the Lanham Act authorizes individuals “dissatisfied with the decision of the [TTAB]” to allow registration of a mark to file a civil action to have the registration canceled and obtain other appropriate relief. 15 U.S.C. § 1071(b)(1). A § 1071(b) action is pursued in lieu of an appeal of the TTAB decision, which must be filed in the U.S. Court of Appeals for the Federal Circuit. *See* § 1071(a) and (b)(1). Therefore, a § 1071(b) action is a new proceeding, and the parties are permitted to submit new evidence beyond the TTAB record. *See Kappos v. Hyatt*, 566 U.S. 431, 439 (2012).

II.

The Lanham Act prohibits registration of a mark that “so resembles a mark registered in the Patent and Trademark Office . . . as to be likely, when used on or in connection with the goods of the applicant, to cause confusion, or to cause mistake or to deceive.” 15 U.S.C. § 1052(d). This standard is met if the new mark is “likely to produce confusion in the minds of consumers about the origin of the goods or services in question.” *Swatch AG v. Beehive Wholesale, LLC*, 739 F.3d 150, 158 (4th Cir. 2014). “To determine if there is a likelihood of confusion between two marks, we consider nine non-exhaustive and non-mandatory factors,” which “serve as a guide rather than ‘a rigid formula’” and “are not all of equal importance” or “relevant in every case.” *Id.* at 158–59 (quoting *George & Co., LLC v. Imagination Entm’t Ltd.*, 575 F.3d 383, 393–94 (4th Cir. 2009)).

The factors, how the district court weighed them, and their status on appeal are represented below:

Factor²	District Court Finding	Status on Appeal
Strength or distinctiveness of the plaintiff’s mark	Favored Combe	Challenged
Similarity of the marks	Favored Combe	Challenged
Similarity of the goods or services the marks identify	Favored Combe	Unchallenged
Similarity of the markholders’ facilities	Favored Combe	Unchallenged

² Drawn from *Pizzeria Uno Corp. v. Temple*, 747 F.2d 1522, 1527 (4th Cir. 1984).

Similarity of the markholders' advertising	Favored Combe	Unchallenged
Defendant's intent	Minimally favored Wolff	Unchallenged
Actual confusion	Favored Combe	Challenged
Quality of the defendant's mark	Did not apply	Unchallenged
Sophistication of the consuming public	Neutral	Challenged

Because the district court entered judgment in favor of Combe following a bench trial, “the standard of review is a deferential one requiring appellate respect for the trial court’s findings.” *Petro Shopping Ctrs., L.P. v. James River Petroleum, Inc.*, 130 F.3d 88, 92 (4th Cir. 1997). “Likelihood of confusion is an inherently factual issue, and [the Court reviews] district court determinations regarding it under a clearly erroneous standard.” *Swatch AG*, 739 F.3d at 155 (internal quotation marks and alteration omitted); *see* Fed. R. Civ. P. 52(a)(6) (“Findings of fact, whether based on oral or other evidence, must not be set aside unless clearly erroneous, and the reviewing court must give due regard to the trial court’s opportunity to judge the witnesses’ credibility.”). A district court’s finding is “‘clearly erroneous’ when although there is evidence to support it, the reviewing court on the entire evidence is left with the definite and firm conviction that a mistake has been committed.” *Anderson v. City of Bessemer City, N.C.*, 470 U.S. 564, 573 (1985).

III.

On appeal, Wolff challenges the district court's determination that Combe had shown a likelihood of confusing Vagisil and Vagisan. Specifically, it asserts the district court erred in assessing four of the factors: strength or distinctiveness of the Vagisil mark, similarity of the two marks, actual confusion, and sophistication of the consumer. We have reviewed the record and conclude that the district court's determinations were not clearly erroneous as to any of these factors.³ Accordingly, we will affirm the judgment of the district court.

A. Strength of the Vagisil Mark

Wolff understandably spends a substantial time arguing that the district court erred in finding that the strength of the mark factor favored Combe given that this is “[t]he first and paramount factor” in the confusion assessment. *Pizzeria Uno Corp. v. Temple*, 747 F.2d 1522, 1527 (4th Cir. 1984). This factor assesses “the degree to which a consumer in the relevant population, upon encountering the mark, would associate the mark with a unique source.” *CareFirst of Md., Inc. v. First Care, P.C.*, 434 F.3d 263, 269 (4th Cir. 2006). “Generally, the stronger the mark, the greater the likelihood that consumers will be confused by competing uses of the mark.” *George & Co.*, 575 F.3d at 393.

A mark's strength consists of two components: conceptual strength and commercial strength. *Id.* Conceptual strength focuses on the mark's own features, measuring its

³ Wolff raises multiple arguments criticizing the district court's findings as to the four challenged factors. We have considered and rejected these arguments even though the opinion does not discuss all of them.

“inherent strength” based on “the linguistic or graphical ‘peculiarity’ of the mark, considered in relation to the product . . . to which the mark attaches.” *CareFirst*, 434 F.3d at 269 (quoting *Perini Corp. v. Perini Constr., Inc.*, 915 F.2d 121, 124 (4th Cir. 1990)). In assessing conceptual strength, marks are divided into four categories, in ascending order of strength: generic (e.g., “bleach” or “copiers”), descriptive (e.g., “5 Minute glue”), suggestive (e.g., “Coppertone[®]”), and arbitrary or fanciful (e.g., “Kodak[®]” or “Exxon[®]”). *George & Co.*, 575 F.3d. at 393–94. Separately, commercial strength “looks at the marketplace and asks if in fact a substantial number of present or prospective customers understand the designation when used in connection with a [product] to refer to a particular” brand. *CareFirst*, 434 F.3d at 269 (internal quotation marks omitted). Assessing commercial strength “entails a rigorous evidentiary standard” weighing six factors: “(1) the plaintiff’s advertising expenditures; (2) consumer studies linking the mark to a source; (3) the plaintiff’s record of sales success; (4) unsolicited media coverage of the plaintiff’s business; (5) attempts to plagiarize the mark; and (6) the length and exclusivity of the plaintiff’s use of the mark.” *George & Co.*, 575 F.3d at 395.

To establish Vagisil’s commercial strength, Combe submitted scores of exhibits documenting its promotional and advertising expenditures (exceeding \$400 million since 1993); its significant sales revenue (exceeding \$1 billion since 1991); its market dominance in both percentage of and numeric rank within its markets (ranging from 50 to 94% of certain products’ market share since the 1990s and several products ranking in the first and second best-selling spots for their markets); its ubiquitous unsolicited media coverage

(ranging from serious print and other media coverage to being featured in a *Saturday Night Live* sketch); and, of course, its ownership of the registered Vagisil mark since 1978.

In addition, Combe submitted a consumer study—the “fame survey”—conducted by its expert witness, Hal Poret. The fame survey had “both an unaided awareness measurement and then an aided awareness measurement.” J.A. 584. The unaided portion asked respondents, who were both male and female, to name brands that came to mind when thinking of a particular product; the aided portion “was measured by asking respondents, one at a time, about eight different terms, one of which was ‘Vagisil,’ and the others were a variety of other [brands] including [a] ‘control,’” Vagizox. J.A. 584. On the unaided portion of the survey, 38.7% of the respondents named Vagisil, which Poret testified was “a very high level, and basically that shows that it’s a very strong mark in consumers’ consciousness in this area, because it came to mind for so many respondents just upon mentioning the category.” J.A. 591. The response rate was even higher for Vagisil’s targeted consumer, female respondents, with “over 50 percent” of women responding “Vagisil” to the unaided portion of the survey. J.A. 591. (That said, Poret also noted that Vagisil’s unaided response rate for men was indicative of commercial strength given that it demonstrated their awareness of a product they would not use.) The results were even higher for the aided portion of the survey, which showed 85% recognizability even after being calibrated to account for false response rates by using the control’s aided recognition rate.

For its part, Wolff proffered evidence to support its contention that Vagisil is a weak mark. For instance, it introduced evidence relating to approximately sixty-six third-party

vaginal-health products that also used the same first four letters as Vagisil. And Wolff's expert witness testified that, in his view, the fame survey was improperly designed. Specifically, he testified the survey was fundamentally flawed because it used only one control when using more might capture a higher false recognition rate. He also opined that the control was ineffective because it used an unusual suffix ("-zox") that was too dissimilar to any actual vaginal-health products.

The district court first found that Vagisil's conceptual strength was "weakened" by evidence of third-party products using the same first four letters. In its view, Vagisil is a suggestive mark, meaning that it has "inherently distinctive" characteristics which do not "describe a function, use, characteristic, size, or intended purpose" of the product, thus requiring "some operation of the imagination to connect [the suggestive mark] with the goods." J.A. 282–83. Despite the comparative strength of a suggestive mark, the district court observed that Vagisil's strength was "weaken[ed]" by evidence that some sixty-six third-party products began with the letters "vagi-." J.A. 284.

What led the district court to find that, overall, the strength factor favored Combe was the significant evidence of Vagisil's commercial strength. In sum, it observed that the evidence demonstrated Vagisil's "considerable commercial success and that a substantial portion of the consuming public understands the VAGISIL mark as signifying a particular source for vaginal care products." J.A. 286. The court viewed the fame survey's results as "substantial and certainly support[ing] a finding that the VAGISIL mark is commercially famous." J.A. 289. It rejected Wolff's challenges to the fame survey, explaining that "the existence of third-party marks alone is insufficient to demonstrate that a mark is

commercially weak” particularly where, as here, most of the marks were “commercially insignificant” because they were either not in use or had “meager sales and advertising figures.” J.A. 290–91. The court also concluded that the criticism of the survey’s design was “unpersuasive” because Wolff had not introduced any empirical evidence to support its expert’s view that the control’s name was strange or that a different or additional control would have affected the results. J.A. 294.

On appeal, Wolff asserts the district court clearly erred in finding the strength factor favored Combe. It does not challenge the district court’s conclusion that Vagisil was a suggestive mark and that its conceptual strength was “weakened” by the existence of a number of third-party marks using the same letters. Instead, Wolff argues the district court should have recognized that the same evidence of third-party products beginning with “vagi-” also demonstrated Vagisil’s commercial weakness and, therefore, proved its overall lack of strength. For instance, Wolff contends consumers are used to distinguishing between multiple products that begin with the letters “vagi-.” And it maintains that Vagicaine products sold at CVS, Rite Aid, Target, and Walmart demonstrate competition in the market with higher sales figures. In addition, it faults the district court for relying on cases involving trade dress disputes as support for why “generic” products like Vagicaine would not demonstrate Vagisil’s commercial weakness. Wolff also reasserts that the fame survey was either inadmissible or should have been afforded little weight due to inherent design flaws. In particular, Wolff contends the survey was poorly designed because it used only one control and its name, Vagizox, was “highly unusual” and dissimilar to other vaginal-health products on the market. Opening Br. 23.

We have considered Wolff's arguments alongside the record and discern nothing clearly erroneous about the district court's finding that this factor favors Combe. As an initial matter, we underscore that Wolff does not take issue with the majority of Combe's evidence of commercial strength. And that evidence is substantial. Over the past four decades, Vagisil has strengthened its status as a leader in its respective markets; the district court ably recounted that evidence, and we will not belabor the point here. None of that evidence is disputed, and its materiality in assessing Vagisil's commercial strength cannot be readily ignored despite Wolff's desire to focus our attention elsewhere.

None of Wolff's arguments about the third-party products that begin with the letters "vagi-" withstands review. The district court plausibly and reasonably explained why it found that evidence did not diminish Vagisil's commercial strength. For example, almost sixty percent of the marks Wolff identified had been abandoned or were found in applications for mark registration and thus had *no* actual commercial presence that could affect Vagisil's strength. J.A. 291 (observing that thirty-nine of the marks were cancelled, expired, or "otherwise 'dead'" and eleven were drawn from mark registration applications and the record lacked any evidence indicating those brands had ever been used in commerce). The handful of marks with any commercial sales had "extremely limited sales, promotion, and recognition by consumers." J.A. 291–92. The district court did not clearly err in concluding that these marks did not mitigate Vagisil's evidence of commercial strength given that the entire reason we have held that third-party use may be relevant to this assessment is if it demonstrates that a conceptually weak mark has not "formed a strong association with a particular source or product in consumer's minds." *Variety Stores, Inc.*

v. Wal-Mart Stores, Inc., 888 F.3d 651, 663 (4th Cir. 2018). The district court did not clearly err in finding that most of these third-party marks have had no opportunity to influence a consumer's mindset, much less her buying choices.

Further, the district court appropriately gave little weight to generic Vagicaine products sold by big-box retailers because consumers do not associate them "as a source-identifying brand," but instead recognize them as the "generic product seek[ing] to imitate VAGISIL's anti-itch cream." J.A. 293. That the court cited cases analyzing a generic product's trade dress to discuss consumer focus when encountering brand name labels versus generic store-specific products is of no moment. The principles drawn from those cases apply equally well in considering whether the evidence regarding Vagicaine sales cut against Vagisil's commercial strength. The district court reasonably concluded "that consumers do not think of 'Vagicaine' as a source-identifying brand; rather, [they] recognize the 'Vagicaine' label only as a signal that the generic product seeks to imitate VAGISIL's anti-itch cream." J.A. 293. As such, this evidence did not "materially diminish" Vagisil's commercial strength. J.A. 293.

As for the fame survey, the district court appropriately considered the evidentiary value of the survey, weighing the testimony of Poret explaining the rationale behind its design and what the results reflected against the testimony of Wolff's counter-expert concerning his view of its alleged flaws. It was within the purview of the district court to undertake that task, and it is not within our purview to second-guess those findings absent some indication of unreasonable fact-finding. *Inwood Labs. Inc. v. Ives Labs., Inc.*, 456 U.S. 844, 855 (1982) (reiterating that in a bench trial the district court has the "unique

opportunity . . . to evaluate the credibility of witnesses and to weigh the evidence,” and its factual findings will be afforded “deference,” and *must* be accepted unless the Court “is left with the definite and firm conviction that a mistake has been committed” (internal quotation marks omitted)). Among other reasons supporting the district court’s finding, we note the district court’s reliance on proof that other pharmaceutical products have used the suffixes “-zox” and “-ox,” lending credibility to the control’s name; the lack of empirical evidence supporting any of the criticisms by Wolff’s expert about the fame survey’s design; and the general alignment of the control’s respondent recognition rate with other actual marks’ “meager recognition rates,” which reinforced the effectiveness of the control for its purpose of curtailing artificially inflated consumer awareness of Vagisil. J.A. 295.

In sum, Wolff’s criticisms are not sufficient to call into question the district court’s finding that, on the whole, Vagisil is a strong and distinctive mark such that this factor favors Combe. We have previously recognized that “if a mark has sufficient commercial strength such that consumers would associate the mark with a unique source, it may be considered strong despite its conceptual weakness.” *Variety Stores*, 888 F.3d at 663 (internal quotation marks and citation omitted). That principle is true here as well.

B. Similarities Between Vagisil and Vagisan

Wolff also contends the district court clearly erred in finding that similarities between the words “Vagisil” and “Vagisan” caused this factor to favor Combe. This factor “focus[es] on whether there exists a similarity in sight, sound, and meaning which would result in confusion.” *George & Co.*, 575 F.3d at 396. Confusion can result from non-identical words; the inquiry is simply whether they are “sufficiently similar” and thus cause

confusion. *See Lone Star Steakhouse & Saloon, Inc. v. Alpha of Va., Inc.*, 43 F.3d 922, 936 (4th Cir. 1995).

In Wolff's view, the district court should not have considered the names as a whole when considering their similarities, but instead should have focused solely on their last three letters to conclude that they were dissimilar and that, consequently, this factor favored it. Specifically, Wolff asserts the first four letters of both words are weak and descriptive, and so should have played no role in analyzing the similarities of the words. It argues that once those letters are set aside, if the district court had focused solely on the words' "dominant" last three letters ("-sil" and "-san," respectively), it would have had to conclude that the words are dissimilar in sight, sound, and meaning.

The legal premise for Wolff's argument lacks support in our caselaw. We have frequently indicated that the similarity inquiry "focus[es] on the dominant portions of the parties' marks," *George & Co.*, 575 F.3d at 396, but we have said so to focus attention on "the non-generic words *in multiword marks*," *Swatch AG*, 739 F.3d at 159 (emphasis added). For example, when considering the similarity of "Pizzeria Uno" and "Taco Uno," we noted that "Pizzeria" and "Taco" were both generic, descriptive words about food, and that "Uno" was the dominant portion of the mark because it was "the word which has meaning and distinction for the consumer." *Pizzeria Uno*, 747 F.2d at 1535. When analyzing the similarity of single-word marks, however, we have *rejected* the "dominant portion" analysis, holding that it would be *inappropriate* to piecemeal such marks. For example, we rejected a party's contention that we should ignore any similarity in the first three letters of "Swatch" and "Swap" when analyzing this factor. *Swatch AG*, 739 F.3d at

159. We reiterated that the proper analysis “compare[s] whole words, not parts.” *Id.* Accordingly, the district court properly considered “Vagisil” and “Vagisan” as whole words.

Moreover, the district court did not clearly err in finding that “Vagisil” and “Vagisan” are “sufficiently similar” for this factor to favor Combe. As the court carefully explained, the two terms “are closely similar in sight and sound,” “consist[ing] of ‘Vagis,’ then a vowel, and then a consonant.” J.A. 297. “Indeed, only two of the seven letters differ between the marks,” and the two unique letters are at the end of each word for purposes of distinguishing their sight and sound. J.A. 297. Likewise, the identical first four letters “connote the vaginal area,” causing the meaning of both words to be similar. J.A. 298.

In short, the court’s determination that this factor favored Combe is well-reasoned and consistent with our precedent. *E.g.*, *Sara Lee Corp. v. Kayser-Roth Corp.*, 81 F.3d 455, 465–66 (4th Cir. 1996) (“L’eggs[®] and Leg Looks[®], although not identical, are perceived similarly by the eye and ear. Whether being written or spoken, L’eggs[®] and the first syllable of Leg Looks[®] are quite similar. . . . [T]here is no substantial difference . . . that would serve to ameliorate any confusion of the[] marks.”).

C. Actual Confusion Between Vagisil and Vagisan

Wolff next challenges the district court’s finding that Vagisil had shown actual confusion between the Vagisil and Vagisan marks. This factor is frequently “persuasive evidence relating to” any likelihood of confusion. *Louis Vuitton Malletier S.A. v. Haute Diggity Dog, LLC*, 507 F.3d 252, 263 (4th Cir. 2007).

Because Vagisil and Vagisan have not been sold on domestic shelves alongside each other, Combe did not rely on evidence of actual consumer confusion in the marketplace, but instead introduced the results of a confusion survey Poret conducted on its behalf. Using the “*Eveready*”⁴ methodology,” Poret surveyed 400 female consumers of vaginal-health products. Under this method, the survey did not mention or refer to Vagisil. Instead, a test group was shown the word “Vagisan” while a control group was shown the control word “Vagipur.” Both words shared the same typography: plain and unadorned block letters lacking any unique features. Respondents were asked several questions concerning what company they believed to be responsible for the product and related matters. Respondents were also given the opportunity to explain their answers. Based on Poret’s tabulation, the survey yielded a statistically significant net confusion rate of 19%, “reflect[ing] that 19% of [survey] respondents experienced confusion that was caused specifically by the marks’ similarity and not by other factors.” J.A. 306. Poret testified about the results of the confusion survey at trial and explained his rationale for its design as well as how he tabulated results.

In response to Combe’s evidence, Wolff’s counter-expert testified to his view that the confusion survey was poorly designed and had been incorrectly tabulated. Apart from that expert testimony critiquing the confusion survey, Wolff did not introduce any evidence demonstrating a lack of actual confusion.

⁴ *Union Carbide Corp. v. Ever-Ready Inc.*, 531 F.2d 366 (7th Cir. 1976), *superseded by rule on other grounds as stated in Scandia Down Corp. v. Euroquilt, Inc.*, 772 F.2d 1423 (7th Cir. 1985).

The district court found that this factor favored Combe because the survey constituted persuasive evidence of actual confusion. It explained why the lack of anecdotal marketplace evidence was understandable and “not meaningful” under the circumstances presented. J.A. 305. Further, the district court found that Poret had used a “reliable approach” that persuasively showed “actual confusion between the parties’ marks.” J.A. 306. And it rejected Wolff’s objections as not materially affecting the results of the survey and therefore, not affecting its finding.

On appeal, Wolff renews its argument that Combe’s confusion survey had design flaws that skewed the results and that Poret compounded those errors in how he tabulated the survey’s results. Specifically, Wolff asserts the survey should not have shown respondents the word “Vagisan” in all-capitals Times New Roman font rather than how it appears on Wolff’s international packaging because that “divorced [it] from [the] marketplace context.” Opening Br. 48. It claims that the control name, “Vagipur,” was ineffective because (1) it did not use an “s” after the prefix to more closely resemble Vagisan; (2) its suffix connoted purity rather than a meaning more in line with Vagisan’s, which is derived from the Latin word *sanitas*, for “health”; and (3) its suffix sounded foreign given it was shared with foreign cities like Kuala Lumpur. Wolff further claims the survey “artificially drove up” the confusion rate based on what counted as confusion with Vagisil in the initial and follow-up answers. Opening Br. 55. Under Wolff’s standard for tabulating the results, the confusion rate would decrease by about half, to 11%, which it asserts is too low to indicate confusion.

The district court did not clearly err in finding that this factor weighed in favor of Combe. At bottom, Wolff’s arguments on appeal ask us to step into the role of the trier-of-fact and reweigh the evidence, a task that is not ours to perform. As we have previously noted, so long as the district court’s factual determinations between conflicting evidence do not leave us with a definite and firm conviction of error, we must not disturb those findings. *June Med. Servs. L. L. C. v. Russo*, ___ U.S. ___, 140 S. Ct. 2103, 2121 (2020) (“In applying this standard to the findings of a district court sitting without a jury, appellate courts must constantly have in mind that their function is not to decide factual issues *de novo*. Where the district court’s account of the evidence is plausible in light of the record viewed in its entirety, the court of appeals may not reverse it even though convinced that had it been sitting as the trier of fact, it would have weighed the evidence differently.” (internal quotation marks, alterations, and citations omitted)). Applied here, we discern no fault in the district court’s reasoning about why the lack of evidence of actual confusion in the marketplace was understandable given that the Vagisil and Vagisan marks had not appeared together in the United States, the relevant market for assessing the existence of actual confusion. Moreover, courts have recognized the role that surveys, such as Poret’s, can serve in assessing actual confusion so long as they use a reliable methodology to elicit relevant evidence. *E.g., 1-800 Contacts, Inc. v. Lens.com, Inc.*, 722 F.3d 1229, 1246 (10th Cir. 2013) (“Surveys can be used as evidence of actual confusion, but their evidentiary value depends on the relevance of the questions asked and the technical adequacy of the survey procedures.”). Any flaws in the survey’s design would ordinarily go to its weight, not its initial admissibility. *Id.*

To the extent Wolff attacks the overarching methodology Poret used, there's no merit to that criticism. The confusion survey followed the "standard and widely accepted" *Eveready* survey methodology. 6 J. Thomas McCarthy, *McCarthy on Trademarks and Unfair Competition* § 32:174 (5th ed. 2021); *see also id.* (describing this method as "especially useful when the senior mark is readily recognized by buyers in the relevant universe" and recognizing that it has been called "the gold standard for fundamental cognitive and marketing reasons"). The district court did not clearly err in finding that this was an appropriate structure for questioning respondents.

Nor has Wolff shown that the district court erred in finding the confusion survey was appropriately designed to assess confusion between the Vagisil and Vagisan marks. Wolff's arguments to the contrary rehash the testimony of its expert witness. But how to view the competing testimony concerning the survey's design—particularly the viability of the control "Vagipur"—was appropriately within the province of the district court to resolve. *See Inwood Labs.*, 456 U.S. at 855; *Helton v. AT & T Inc.*, 709 F.3d 343, 350 (4th Cir. 2013). The district court explained its reasoning for why "Vagipur" "was adequate to measure survey 'noise,'" including that its first four letters were identical and thus also "connoted the vaginal area of a woman's body"; those letters were followed by the same number of letters in the same consonant-vowel-consonant order; and the control did not use the letter "s" because that was part of what made Vagisan "confusingly similar to" Vagisil. J.A. 309–10. Wolff has not shown that this reasoning was clearly erroneous.

The evidence also does not compel the conclusion that the district court erred in evaluating the survey results. The district court explained that the approach Wolff's expert

advocated “would improperly double-count survey ‘noise’” through the use of both an effective control *and* sifting through the explanations for a respondent’s confusion. J.A. 311. Because the control provided an objective—and therefore better—measure for “estimating the number of respondents that will name VAGISIL for reasons other than the similarity between VAGISAN and VAGISIL,” it was not necessary to use the “scientifically inferior” method of relying on a lay person’s attempt to explain their response and a survey administrator’s judgment as to whether that indicated actual confusion. J.A. 311.

Because Wolff has not shown that the district court’s assessment of the actual confusion factor was clearly erroneous, we reject it as a proper ground for revisiting the district court’s judgment.

D. Neutrality of Consumer Sophistication

Lastly, Wolff contends the district court clearly erred in finding the sophistication of the consumer to be a neutral factor because the record contained evidence that consumers selected vaginal-health products with a high degree of care. In its view, this testimony meant that the factor should weigh in its favor because informed consumers would be less likely to confuse the marks. We disagree.

In describing this factor, we have previously recognized that it plays a “key” role in the assessment only “when the relevant market is not the public at-large.” *Sara Lee Corp.*, 81 F.3d at 467. “[I]n a market with extremely sophisticated buyers, the likelihood of consumer confusion cannot be presumed on the basis of the similarity in trade name alone, particularly without the benefit of trial.” *Perini Corp.*, 915 F.2d at 128. The relevant market

must consist of a targeted subset that is either “sophisticated in the use of—or possesses an expertise regarding”—the product. *Sara Lee Corp.*, 81 F.3d at 467. Thus, for example, in a trademark infringement suit between two hosiery makers, we held it was not necessary to consider this factor because there was no evidence “that persons who buy pantyhose are any more sophisticated about that product than those who comprise the market for other ordinary retail goods.” *Id.*

Applied here, there’s no indication in the record that consumers are “extremely sophisticated buyers” that would warrant this factor having any sort of a “key” role in the analysis. Rather, as was true of hosiery, vaginal-health products are marketed broadly to female consumers as a whole. And while the record contained evidence that some women were well-informed about these products given their nature, there was also evidence that some women approached purchases in a manner consistent with lack of sophistication. That is, some shoppers exhibited “shy to buy” characteristics of “get[ing] in and out of the store” hurriedly, putting “products in their bag real[ly] quickly and put[ting] potato chips on it and run[nin]g out of the store so people don’t see it” because the products may elicit embarrassment or shame in their consumers. J.A. 466–67. After a bench trial, the district court weighed this evidence and determined that the factor was neutral and favored neither Wolff nor Combe. Nothing in this record leaves us with the definite or firm conviction that the district court erred in so finding. Accordingly, we reject this argument as well.

IV.

Given that the district court did not clearly err as to any of its individual assessments of the challenged factors, we have no basis for reconsidering its overarching assessment that the factors demonstrated a reasonable likelihood of confusing the Vagisil and Vagisan marks. The court did not err in entering judgment in favor of Combe and ordering appropriate relief. For these reasons, the judgment of the district court is

AFFIRMED.