

# Stanford **Technology** Law Review

## Consistency of Confusion? A Fifteen-Year Revisiting of Barton Beebe's Empirical Analysis of Multifactor Tests for Trademark Infringement

KEVIN BLUM, ARIEL FOX, CHRISTINA J. HAYES, AND JAMES (HANJUN) XU\*

CITE AS: 2010 STAN. TECH. L. REV. 3

<http://stlr.stanford.edu/pdf/blum-consistency-of-confusion.pdf>

---

\* All authors are J.D., Harvard Law School, 2009. Initially prepared for Professor William Fisher's Spring 2009 Advanced Intellectual Property Class. The authors wish to thank Professors William Fisher, Barton Beebe, and James Greiner for their invaluable advice and support.

**TABLE OF CONTENTS**

INTRODUCTION ..... 3

I. METHODOLOGY AND SUMMARY STATISTICS..... 3

*Methodology*..... 3

*Summary Statistics* ..... 4

II. INTERFACTOR AND FACTOR ANALYSIS..... 11

*Core Factors*..... 11

*Correlation Tables* ..... 14

*Stampeding*..... 20

*Historical Trends*..... 24

*Factor 1: The Strength of the Mark*..... 24

*Factor 2: The Similarity of the Marks*..... 27

*Factor 3: The Proximity of the Products*..... 28

*Factor 4: Bridging the Gap*..... 28

*Factor 5: Actual Confusion* ..... 30

*Factor 6: Intent* ..... 31

*Factor 7: The Quality of Defendant’s Mark*..... 32

*Factor 8: The Sophistication of Consumers*..... 34

PART III: OTHER CONSIDERATIONS ..... 34

*Competition* ..... 34

*“Any One Factor May Be Dispositive”* ..... 35

*Party Names* ..... 35

*Additional Factors* ..... 35

CONCLUSION..... 36

APPENDIX ..... 37

## INTRODUCTION

¶1 In a typical case of trademark infringement, a plaintiff must show, among other things, that potential consumers would be confused as to the source of a good or service, due to the defendant’s use of a mark (or in the case of trade dress infringement, the same or similar product packaging, design, labeling, etc.) that creates confusion as to source, sponsorship, or affiliation. This inquiry into the likelihood of confusion is most often governed by multifactor tests, the most prominent of which is the test articulated in *Polaroid Corp. v. Polarad Electronics Corp.*<sup>1</sup> by the U.S. Court of Appeals for the Second Circuit. The Second Circuit is of particular interest to those who study trademark law as it generates the greatest number of trademark infringement opinions.<sup>2</sup> In the *Polaroid* test, eight factors are considered, including: the strength of the original user’s mark, similarity of the marks, proximity of the products, the likelihood that the original user would enter into the alleged infringer’s market (thus, “bridging the gap” between the two markets), evidence of actual confusion, the alleged infringer’s intent, the relative quality of the products bearing the marks in question, and sophistication of the consumers of the products.

¶2 Professor Barton Beebe, in his empirical analysis of multifactor tests for trademark infringement nationwide, found that judges “employ ‘fast and frugal’ heuristics to short-circuit the multifactor analysis” and challenged the conventional wisdom that no single factor in the multifactor test is dispositive.<sup>3</sup> However, the limited temporal scope of his inquiry (2000–2004) raises the question of whether his results are typical for trademark cases or just a five-year anomaly. This paper attempts both to determine whether Beebe’s findings hold true over a longer, fifteen-year period and to discover any new information that sheds light on how judges decide trademark cases. Part I.A describes our methodology while Part I.B provides background statistics on our data set. Part II presents analysis of factor correlation, outcome classification trees, stampeding, and each individual factor, including comparison with Beebe’s results, as well as description of historical trends in our data. Part III analyzes the effect of other considerations related to the multifactor test—whether or not the products are in competition, whether the party names are the trademarks at issue, whether the court indicates one factor can be dispositive of the outcome, and whether or not the court considers factors in addition to those listed in the Second Circuit’s *Polaroid* test.

## I. METHODOLOGY AND SUMMARY STATISTICS

### *Methodology*

#### *Determining the Sample*

¶3 The goal of testing Beebe’s results over a longer period of time needed to be balanced with the time and labor constraints of four students working over the course of one semester. This balance was struck by analyzing cases from the U.S. District Court for the Southern District of New York decided between January 1, 1994 and December 31, 2008. The U.S. District Court for S.D.N.Y. was chosen because it is governed by the Second Circuit, which, as mentioned above, is the leading court for deciding trademark infringement. This court was further singled out for analysis because first, most trademark infringement opinions at the trial level are generated there,<sup>4</sup> and second, as a trial court governed by the Second Circuit, we could be relatively certain that its judges would apply the established *Polaroid* multifactor test. Accordingly, we determined our initial list of cases by adding the word “Polaroid” to the search terms Beebe used in his paper;<sup>5</sup> because the *Polaroid* test is the multifactor test that we chose to analyze. The occasional opinions that ignored the test were

---

<sup>1</sup> 287 F.2d 492, 495 (2d Cir. 1961).

<sup>2</sup> Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 94 CAL. L. REV. 1581, 1594-96 (2006).

<sup>3</sup> *Id.* at 1581-82.

<sup>4</sup> Beebe, *supra* note 1, at 1595 n. 64.

<sup>5</sup> Beebe, *supra* note 1, at 1640 n. 262. Our exact search terms were: (“trademark infringement” & confus!) or ((trademark mark) & “likelihood of confusion”) & Polaroid.”

of questionable worth to our project. After generating this initial list, we checked for duplicate cases and removed them. We then screened the case list for “substantial use,” *i.e.*, “any use beyond the mere citation without analysis of the test.”<sup>6</sup> After removing the cases without substantial use, the list contained 224 cases.

### *Coding*

¶4 Before coding the cases, we formulated a list of variables and tested that list on five substantial use trademark cases from the U.S. District Court for the Eastern District of New York. We then met to discuss why we coded those cases the way we did so as to ensure that all cases were analyzed in the same way. From that meeting, we created a final list of variables, explanation of the variables, and rules for coding.<sup>7</sup>

¶5 Cases were randomly assigned for coding among the four group members.<sup>8</sup> All cases were coded by two people so as to ensure accuracy. A list of double-coded cases was generated, with discrepancies between the two coders’ versions highlighted. A third person then acted as a tie-breaker for all discrepancies between the two coders’ versions.<sup>9</sup> The third person had to choose between the two options presented to him; he could not code a third choice if he felt that both of the original coders were wrong. From this final list, the following categories of cases were excluded due to the likelihood that they would not provide accurate information about how judges apply the multifactor test: (1) opinions that discuss likelihood of confusion in the counterfeiting context; (2) opinions where trademark infringement appears alongside a breach of a franchising, licensing, or distribution agreement; (3) opinions where the non-moving party failed to appear; (4) opinions on a motion to dismiss; and (5) opinions where the outcome of the *Polaroid* test was reversed on appeal. After removing these cases, the final coded case list contained 206 unique opinions.

### *Analysis*

¶6 Seven different types of analysis were performed: (1) creation of outcome classification trees using Beebe’s methodology;<sup>10</sup> (2) calculation of plaintiff win rates; (3) analysis of the interactions between the strength of the mark, commercial strength, inherent strength, and the *Abercrombie* factors using (a) simple descriptive statistics and (b) a two-tailed Probit regression with a 95% confidence interval; (4) creation of correlation tables using Stata’s pairwise correlation function at a 95% confidence interval; (5) calculation of stampeding scores using Beebe’s methodology; (6) analysis of historical trends using basic descriptive statistics; and (7) simple descriptive analyses of the factors as well as additional non-*Polaroid* factors. Analysis was performed on both the entire data set and a data set including only preliminary injunction and bench trial opinions (the “PI/BT” data set). The full data set may be found at [http://spreadsheets.google.com/pub?key=r9UGu9rqUFs\\_uUg4g4OVSrw](http://spreadsheets.google.com/pub?key=r9UGu9rqUFs_uUg4g4OVSrw).

### *Summary Statistics*

¶7 The most common procedural posture for our sample of 206 cases was an opinion on a preliminary injunction (44.7%). This is consistent with Beebe’s observation that preliminary injunctions are the primary forum for the adjudication of trademark infringement claims.<sup>11</sup> Table A sets out the number of opinions by posture and percentage of the full sample that they represent.

¶8 Table A also includes information about plaintiff and defendant win rates. Plaintiff<sup>12</sup> win rates are the proportion of opinions that found a likelihood of confusion, and defendant win rates are the proportion of

---

<sup>6</sup> *Id.* at 1650.

<sup>7</sup> See Appendix, Table 1, for rules and explanations of variables.

<sup>8</sup> Cases were randomized in the following way: The cases were randomized using the “list randomizer” found at <http://www.random.org/lists>. The cases were then numbered based on the random order in which they appeared. Person A was assigned cases 1-112, Person B was assigned cases 81-192, Person C was assigned cases 1-50 and 163-224, and Person D was assigned cases 51-80, 113-162, and 193-224. The identities of Persons A-D were randomly determined by running the authors’ last names through the aforementioned “list randomizer.”

<sup>9</sup> Person A re-coded cases 127-166 and 199-214, Person B re-coded cases 25-64, 193-198, and 215-224, Person C re-coded cases 65-106 and 113-126, and Person D re-coded cases 1-24, 107-112, and 167-192.

<sup>10</sup> Beebe, *supra* note 1, at 1603-14.

<sup>11</sup> Beebe, *supra* note 1, at 1595.

<sup>12</sup> Plaintiff, as used in this section and elsewhere in this note, does not necessarily mean the actual litigation plaintiff; it is merely used

opinions that did not find a likelihood of confusion.<sup>13</sup> Our sample includes only litigated federal trademark infringement cases that produced written opinions returned by our search methods in the Westlaw and Lexis databases, and is exclusive of opinions where the court did not make substantial use of the multifactor test. Since, as noted above, we also excluded opinions meeting certain criteria from our data set, we are cautious about the scope of the inferences we can draw from the win rate data.

¶9 We found that slightly less than half of the opinions in our sample resulted in a finding of a likelihood of confusion (a 41.3% plaintiff win rate). This is somewhat higher than the plaintiff win rate observed by Beebe in his 2000–2004 sample of Southern District of New York opinions (36%), but still lower than the overall plaintiff win rate observed by Beebe nationally over the same period of time (51%). We observed a higher win rate in preliminary injunction and bench trial opinions (48.9%), consistent with both Beebe’s findings and the “fifty percent hypothesis.”<sup>14</sup>

¶10 Unsurprisingly, the highest plaintiff win rate (61.5%)<sup>15</sup> was found in summary judgment opinions where the movant was the proponent of the likelihood of confusion analysis and the opponent did not move for cross-summary judgment. This likely reflects, in part, the litigation strategy of attempting to dispose of stronger claims on summary judgment. As noted by Beebe, it may also reflect the fact that a judge is more likely to produce a written opinion when he or she grants a summary judgment motion rather than when he or she denies it.<sup>16</sup> Conversely, and also unsurprisingly, the highest defendant win rate (81.5%) was found in summary judgment opinions where the movant was the opponent of the likelihood of confusion analysis and the proponent did not move for cross-summary judgment.

---

as shorthand for the proponent of a finding of likelihood of confusion. The converse is true of the term “defendant,” which is used as shorthand for the opponent of a finding of likelihood of confusion. In the case of a declaratory judgment action, the defendant would be considered a plaintiff in the plaintiff win rate calculation, and the plaintiff, proponent of a finding of no likelihood of confusion, would be considered a defendant in the defendant win rate calculation.

<sup>13</sup> Opinions where the judge did not make a determination of the outcome of the likelihood of confusion analysis are not included in the plaintiff or defendant win rate calculations.

<sup>14</sup> See George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 6-30 (setting out the hypothesis that we should observe a 50% plaintiff win rate regardless of the substantive standard of liability, as disputes in which the outcome is more easily and accurately predicted by parties are more likely to be settled than litigated).

<sup>15</sup> Excluding the temporary restraining order opinion and the magistrate judge report and recommendation categories, which had only one opinion each.

<sup>16</sup> Beebe, *supra* note 1, at 1598.

Table A

**PLAINTIFF WIN RATE BY PROCEDURAL POSTURE**  
**(PW = plaintiff multifactor test win rate, DW = defendant multifactor test win rate)**

	All Opinions		Dispositive Opinions		Preliminary Injunction		Bench Trial		SJ Plaintiff	
	N	PW	N	PW	N	PW	N	PW	N	PW
	206	0.413	139	0.489	92	0.554	48	0.354	13	0.615
% of Total	100.0		67.5		44.7		23.3		6.3	

	SJ Defendant		SJ Cross		TRO		Magistrate RR	
	N	DW	N	DW	N	PW	N	PW
	27	0.815	24	0.292	1	1.000	1	0.000
% of Total	13.1		11.7		0.5		0.5	

SJ = Summary Judgment

TRO = Temporary Restraining Order Opinion

Magistrate RR = Magistrate Judge Report and Recommendation

**Table B**

PLAINTIFF WIN RATE AND STAMPEDE SCORE BY FACTOR OUTCOME IN  
 139 PRELIMINARY INJUNCTION AND BENCH TRIAL OPINIONS  
 (LoC = likelihood of confusion, PW = plaintiff multifactor test win rate  
 SS = stampede score, SD = standard deviation of SS)

Factor	Outcome	Distribution by Outcome		PW	Stampede Score			
		N	%		SS	SD	Skew	Kurtosis
1 Strength of the mark	Favors LoC	88	63.3	0.739	0.385	0.437	-0.781	0.005
	Disfavors LoC	39	28.1	0.000	-0.487	0.264	0.176	-0.555
	Other	12	8.6	0.250	-0.198	0.343	1.057	0.493
2 Similarity of the marks	Favors LoC	86	61.9	0.791	0.387	0.444	-0.741	-0.366
	Disfavors LoC	48	34.5	0.000	-0.388	0.336	0.334	-0.356
	Other	5	3.6	0.000	-0.425	0.259	1.447	1.931
3 Proximity of the products	Favors LoC	92	66.2	0.685	0.370	0.428	-0.514	-0.720
	Disfavors LoC	33	23.7	0.061	-0.508	0.269	0.566	0.160
	Other	14	10.1	0.214	-0.339	0.358	0.282	-1.090
4 Likelihood of bridging the gap	Favors LoC	47	33.8	0.723	0.468	0.391	-0.392	-0.837
	Disfavors LoC	30	21.6	0.100	-0.513	0.291	0.641	-0.015
	Other	62	44.6	0.500	0.095	0.497	-0.099	-1.388
5 Evidence of actual confusion	Favors LoC	39	28.1	0.923	0.587	0.376	-1.427	2.108
	Disfavors LoC	59	42.4	0.136	-0.294	0.414	0.555	-0.461
	Other	41	29.5	0.585	0.171	0.455	-0.290	-1.035
6 Bad faith	Favors LoC	43	30.9	0.884	0.613	0.340	-1.543	3.443
	Disfavors LoC	72	51.8	0.208	-0.255	0.404	0.441	-0.430
	Other	24	17.3	0.625	0.188	0.473	-0.681	-1.020
7 Quality of the products	Favors LoC	34	24.5	0.706	0.574	0.419	-0.914	-0.349
	Disfavors LoC	29	20.9	0.310	-0.263	0.496	0.591	-0.886
	Other	76	54.7	0.461	0.008	0.482	-0.087	-1.316
8 Consumer sophistication	Favors LoC	41	29.5	0.805	0.570	0.363	-0.803	-0.159
	Disfavors LoC	60	43.2	0.233	-0.302	0.423	0.643	-0.501
	Other	38	27.3	0.553	0.191	0.443	-0.399	-1.094
	Total opinions	139		0.489	0.090	0.553	-0.068	-1.248
	LoC found	68	48.9		0.570	0.267	-0.359	-0.126
	LoC not found	71	51.1		-0.370	0.311	0.358	-0.156

Table C

FACTOR OUTCOME AND STAMPEDE SCORE BY TEST OUTCOME IN  
 139 PRELIMINARY INJUNCTION AND BENCH TRIAL OPINIONS  
 (LoC = likelihood of confusion, PW = plaintiff multifactor test win rate  
 SS = stampede score, SD = standard deviation of SS)

Factor	Outcome	Test Outcome							
		Opinions In Which LoC Found				Opinions In Which LoC Not Found			
		N	%	SS	SD	N	%	SS	SD
1 Strength of the mark	Favors LoC	65	95.6	0.583	0.262	23	32.4	-0.174	0.337
	Disfavors LoC	0	0.0	---	---	39	54.9	-0.487	0.264
	Other	3	4.4	0.292	0.260	9	12.7	-0.361	0.159
2 Similarity of the marks	Favors LoC	68	100.0	0.570	0.267	18	25.4	-0.306	0.251
	Disfavors LoC	0	0.0	---	---	48	67.6	-0.388	0.336
	Other	0	0.0	---	---	5	7.0	-0.425	0.259
3 Proximity of the products	Favors LoC	63	92.6	0.609	0.233	29	40.8	-0.151	0.255
	Disfavors LoC	2	2.9	0.000	0.177	31	43.7	-0.540	0.240
	Other	3	4.4	0.125	0.125	11	15.5	-0.466	0.286
4 Likelihood of bridging the gap	Favors LoC	34	50.0	0.654	0.256	13	18.3	-0.019	0.227
	Disfavors LoC	3	4.4	0.000	0.125	27	38.0	-0.569	0.244
	Other	31	45.6	0.532	0.209	31	43.7	-0.343	0.252
5 Evidence of actual confusion	Favors LoC	36	52.9	0.660	0.273	3	4.2	-0.292	0.361
	Disfavors LoC	8	11.8	0.406	0.160	51	71.8	-0.404	0.323
	Other	24	35.3	0.490	0.239	17	23.9	-0.279	0.260
6 Bad faith	Favors LoC	38	55.9	0.694	0.230	5	7.0	0.000	0.442
	Disfavors LoC	15	22.1	0.317	0.245	57	80.3	-0.406	0.285
	Other	15	22.1	0.508	0.160	9	12.7	-0.347	0.292
7 Quality of the products	Favors LoC	24	35.3	0.807	0.173	10	14.1	0.013	0.266
	Disfavors LoC	9	13.2	0.389	0.211	20	28.2	-0.556	0.228
	Other	35	51.5	0.454	0.217	41	57.7	-0.372	0.273
8 Consumer sophistication	Favors LoC	33	48.5	0.708	0.225	8	11.3	0.000	0.250
	Disfavors LoC	14	20.6	0.277	0.256	46	64.8	-0.478	0.283
	Other	21	30.9	0.548	0.150	17	23.9	-0.250	0.230
	LoC found	68		0.570	0.267				
	LoC not found					71		-0.370	0.311

Table D

PLAINTIFF WIN RATE AND STAMPEDE SCORE BY FACTOR OUTCOME IN 119 PRELIMINARY INJUNCTION AND BENCH TRIAL OPINIONS, COMPETING AND NON-COMPETING GOODS/SERVICES  
(LoC = likelihood of confusion, PW = plaintiff multifactor test win rate  
SS = stampede score, SD = standard deviation of SS)

Factor	Outcome	Distribution by Outcome		PW	Stampede Score			
		N	%		SS	SD	Skew	Kurtosis
1 Strength of the mark	Favors LoC	77	64.7	0.727	0.380	0.44907	-0.751	-0.0884
	Disfavors LoC	34	28.6	0.000	-0.460	0.26422	0.03663	-0.4596
	Other	8	6.7	0.250	-0.266	0.32347	1.17168	1.24652
2 Similarity of the marks	Favors LoC	74	62.2	0.784	0.380	0.460	-0.694	-0.516
	Disfavors LoC	40	33.6	0.000	-0.363	0.345	0.275	-0.409
	Other	5	4.2	0.000	-0.425	0.259	1.447	1.931
3 Proximity of the products	Favors LoC	83	69.7	0.651	0.346	0.441	-0.393	-0.904
	Disfavors LoC	28	23.5	0.036	-0.563	0.222	0.107	-0.765
	Other	8	6.7	0.375	-0.188	0.341	0.141	-2.480
4 Likelihood of bridging the gap	Favors LoC	40	33.6	0.675	0.453	0.418	-0.297	-1.124
	Disfavors LoC	28	23.5	0.107	-0.522	0.266	0.575	0.206
	Other	51	42.9	0.549	0.157	0.478	-0.258	-1.313
5 Evidence of actual confusion	Favors LoC	33	27.7	0.909	0.580	0.406	-1.318	1.417
	Disfavors LoC	52	43.7	0.154	-0.262	0.422	0.460	-0.649
	Other	34	28.6	0.588	0.176	0.465	-0.271	-1.159
6 Bad faith	Favors LoC	36	30.3	0.889	0.642	0.328	-1.744	5.159
	Disfavors LoC	62	52.1	0.194	-0.258	0.385	0.485	-0.052
	Other	21	17.6	0.667	0.208	0.477	-0.859	-0.717
7 Quality of the products	Favors LoC	31	26.1	0.710	0.573	0.427	-0.922	-0.338
	Disfavors LoC	24	20.2	0.292	-0.250	0.476	0.710	-0.480
	Other	64	53.8	0.453	-0.004	0.486	-0.096	-1.286
8 Consumer sophistication	Favors LoC	36	30.3	0.806	0.587	0.377	-0.917	-0.061
	Disfavors LoC	50	42.0	0.200	-0.325	0.382	0.568	-0.448
	Other	33	27.7	0.576	0.201	0.436	-0.582	-0.902
	Total opinions	119		0.487	0.097	0.553	-0.046	-1.239
	Competing	73	61.3	0.616	0.310	0.486	-0.459	-0.808
	Non-competing	46	38.7	0.283	-0.242	0.484	0.757	-0.310

Table E

FACTOR OUTCOME AND STAMPEDE SCORE BY COMPETITION IN 119 PRELIMINARY INJUNCTION  
AND BENCH TRIAL OPINIONS, COMPETING AND NON-COMPETING GOODS/SERVICES  
(LoC = likelihood of confusion, PW = plaintiff multifactor test win rate  
SS = stampede score, SD = standard deviation of SS)

Factor	Outcome	Competition of Products									
		Opinions Involving Competing Products					Opinions Involving Non-Competing Products				
		N	%	PW	SS	SD	N	%	PW	SS	SD
1 Strength of the mark	Favors LoC	53	72.6	0.811	0.533	0.334	24	52.2	0.542	0.042	0.490
	Disfavors LoC	15	20.5	0.000	-0.333	0.257	19	41.3	0.000	-0.559	0.230
	Other	5	6.8	0.400	-0.125	0.331	3	6.5	0.000	-0.500	0.125
2 Similarity of the marks	Favors LoC	50	68.5	0.900	0.553	0.334	24	52.2	0.542	0.021	0.484
	Disfavors LoC	22	30.1	0.000	-0.227	0.317	18	39.1	0.000	-0.528	0.311
	Other	1	1.4	0.000	0.000	---	4	8.7	0.000	-0.531	0.120
3 Proximity of the products	Favors LoC	67	91.8	0.657	0.368	0.450	16	34.8	0.625	0.258	0.407
	Disfavors LoC	4	5.5	0.000	-0.594	0.188	24	52.2	0.042	-0.557	0.230
	Other	2	2.7	0.500	0.188	0.088	6	13.0	0.333	-0.313	0.293
4 Likelihood of bridging the gap	Favors LoC	26	35.6	0.692	0.543	0.409	14	30.4	0.643	0.286	0.394
	Disfavors LoC	1	1.4	0.000	-0.750	---	27	58.7	0.111	-0.514	0.267
	Other	46	63.0	0.587	0.201	0.464	5	10.9	0.200	-0.250	0.442
5 Evidence of actual confusion	Favors LoC	23	31.5	0.957	0.739	0.238	10	21.7	0.800	0.213	0.483
	Disfavors LoC	27	37.0	0.259	-0.065	0.412	25	54.3	0.040	-0.475	0.323
	Other	23	31.5	0.696	0.321	0.386	11	23.9	0.364	-0.125	0.487
6 Bad faith	Favors LoC	34	46.6	0.912	0.673	0.250	2	4.3	0.500	0.125	1.061
	Disfavors LoC	27	37.0	0.148	-0.181	0.295	35	76.1	0.229	-0.318	0.437
	Other	12	16.4	0.833	0.385	0.382	9	19.6	0.444	-0.028	0.507
7 Quality of the products	Favors LoC	23	31.5	0.783	0.685	0.333	8	17.4	0.500	0.250	0.522
	Disfavors LoC	11	15.1	0.364	-0.068	0.492	13	28.3	0.231	-0.404	0.421
	Other	39	53.4	0.590	0.196	0.423	25	54.3	0.240	-0.315	0.413
8 Consumer sophistication	Favors LoC	30	41.1	0.800	0.608	0.375	6	13.0	0.833	0.479	0.399
	Disfavors LoC	18	24.7	0.222	-0.160	0.406	32	69.6	0.188	-0.418	0.340
	Other	25	34.2	0.680	0.290	0.378	8	17.4	0.250	-0.078	0.513
	Competing	73		0.616	0.310	0.486					
	Non-competing						46		0.283	-0.242	0.484

## II. INTERFACTOR AND FACTOR ANALYSIS

### *Core Factors*

¶11 Beebe’s work demonstrates the importance of core factors across the federal circuit courts—knowing the results of only a few factors is typically enough to predict the test’s outcome.<sup>17</sup> People—in particular judges—may come to their decisions after considering “only a small number of cues,” and according to the “core attributes heuristic model” “stop acquiring and analyzing information once the last in their set of the most important, determinant attributes has been acquired and analyzed.”<sup>18</sup> Beebe argued it was particularly intelligent for judges to utilize the core attributes heuristic in the likelihood of confusion context, as it meant judges were “taking advance of the ecology of the multifactor test and, in particular, the redundancy of many of its factors to reach a conclusion efficiently.”<sup>19</sup> Our work also indicates that certain factors have more importance than others.

### *Outcome Classification Trees*

¶12 Our findings are consistent with Beebe’s—not only in supporting a core attributes heuristic, but in supporting the importance of certain factors specifically. We created outcome classification trees at one, two, and five levels, as Beebe did; however, working independently without using his questions as guides, we still saw that, over the last 15 years in the SDNY, our trees could fairly accurately predict the outcome of the test knowing only a handful of factors.<sup>20</sup> Each outcome classification tree is the result of asking, at each node, if a certain factor favors or disfavors confusion, and then seeing if the cases split along those lines match the predicted outcome.<sup>21</sup> The most predictive trees are below, with others in the Appendix.

---

<sup>17</sup> Beebe, *supra* note 1, at 1603.

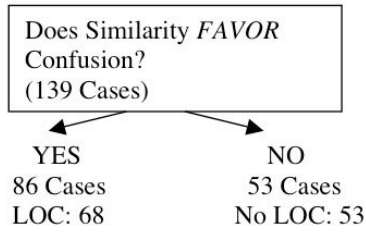
<sup>18</sup> *Id.* at 1601-02.

<sup>19</sup> *Id.* at 1602-03.

<sup>20</sup> These outcome classification trees were done with the 139 Preliminary Injunction and Bench Trial cases only to ensure that there was either a win or a loss. They were also created with subsets of the data—competing products and non-competing products. There was no significant difference in the results, except that the proximity factor became fairly useless as a predictor for products in competition, as proximity favored likelihood of confusion in almost all of the competing products.

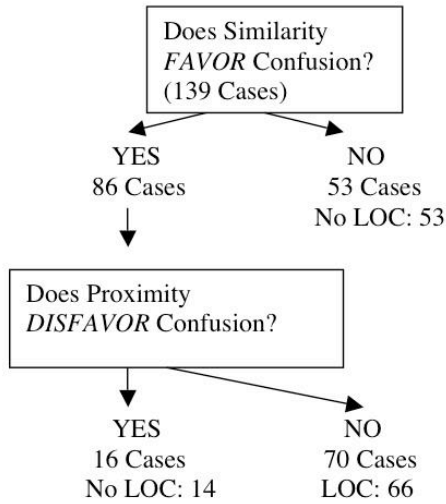
<sup>21</sup> As an example, when asking “Does Similarity *favor* Confusion?” the predicted outcome for any “Yes” cases is that there will be a Likelihood of Confusion finding, and the predicted outcome for a “No” case (which includes findings that similarity disfavors confusion as well as similarity “other” findings), is that there will be No Likelihood of Confusion.

Figure I



**With 1 Node:**  
121 accurate/  
139 total  
87.0% accurate  
100% classified

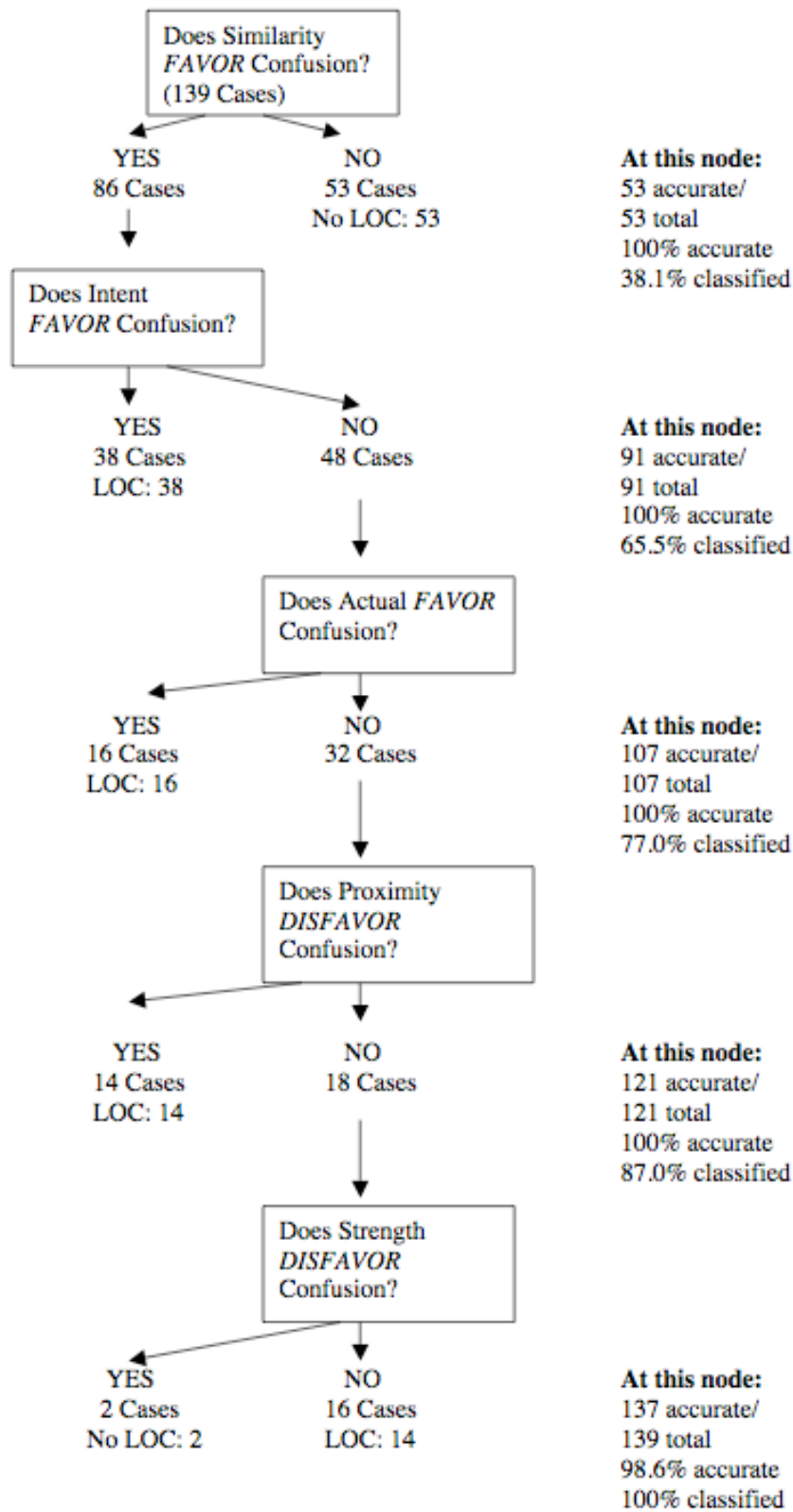
Figure II



**At this node:**  
53 accurate/  
53 total  
100% accurate  
38.1% classified

**At this node:**  
133 accurate/  
139 total  
95.7% accurate  
100% classified

Figure III



¶13 Starting first at the one-node level, asking if Similarity *favours* Confusion, and only that, we could accurately predict the outcome for 87% of the cases—121 out of 139. This is the same question Beebe found most predictive for his one node analysis, at about the same level.<sup>22</sup> The next most useful questions at the first level were if Similarity *disfavours* Confusion (83.5% or 116 cases accurately classified), and if Strength *favours* Confusion (81.3% or 113 cases accurately classified).

¶14 In the two-node classification tree, asking if Similarity *favours* Confusion and then if Proximity *disfavours* Confusion accurately classified 133, or 95.7%, of the 139 cases. These again turned out to be the questions Beebe found most useful for his two-node scheme, classifying 95.8% of his decisions.<sup>23</sup> Asking whether Similarity *favours* Confusion and if Strength *disfavours* Confusion, in either order, correctly predicted the outcome for 131 or 94.2% of cases.

¶15 The best five-node tree asks if Similarity *favours* Confusion, then if Intent *favours* Confusion, then if Actual Confusion *favours* Confusion, then if Proximity *disfavours* Confusion, then if Strength *disfavours* Confusion. At this point, 137 cases, or 98.6%, out of 139, are accurately classified. That said, doing a different five-factor test beginning with if Strength *disfavours* Confusion (with a different questions for the other four factors/nodes), or if Similarity *disfavours* Confusion (with the same factor order and question for all nodes except the last strength node), gives us 136 cases classified, only one less than the best tree.

¶16 The differences between the best classification outcome trees and the second- and third-best are often quite small and perhaps statistically insignificant given the size of our data set. This lends support to the idea that a number of core factors are considered but may not be considered in any coherent order.<sup>24</sup> Nonetheless, our results, like Beebe’s, indicate that “at least some aspects of the multifactor analysis” “are noncompensatory in nature,” meaning that a party cannot make up for losing them.<sup>25</sup> Certain factor outcomes are weighted so strongly as to outweigh the combined weights of all other factor outcomes.”<sup>26</sup> Similarity seems to be the most important factor—and if similarity does not favor confusion, the plaintiff is unlikely to win regardless of the other factors. Strength is also an important factor to win, and intent is important to win if it favors confusion.

#### *Correlation Tables*

¶17 Two correlation tables are included, one for our “all cases” dataset (Table F), and one for only those cases that were decided at the preliminary injunction or bench trial (PI/BT) level (this dataset primarily excludes summary judgment decisions) (Table G).

¶18 The correlation tables show the pairwise correlation values among the eight *Polaroid* factors, two additional factors that we chose to investigate (use of plaintiff’s/defendant’s name in the marks, and whether the products are competing), and the ultimate finding as to whether there is a likelihood of confusion. Each of the factors is further broken down into elements, based upon whether the court found that analysis of the factor supports a likelihood of confusion (supports LoC), militates against one (against LoC), or if the analysis is inconclusive, neutral, or irrelevant (other).

¶19 The intersecting cell between a row and a column of elements reveals a correlation value between -1 and 1 (in which case the two elements are perfectly inversely correlated or perfectly positively correlated, respectively). A star (\*) next to the correlation values signifies that the value is significant to the 0.05 level. In the all cases dataset, we analyzed 206 observations, and in the PI/BT dataset, we analyzed 139 observations. In the PI/BT dataset, all cases concluded that either there was a likelihood of confusion or not as to the ultimate finding—there were no cases that found outstanding issues of fact. As such, the outstanding issues of fact rows and columns for that dataset were omitted.

---

<sup>22</sup> Beebe, *supra* note 1, at 1605 fig.1.

<sup>23</sup> *Id.* at 1605 fig.2.

<sup>24</sup> Beebe cautions that his results do not reflect the “*process* of reasoning the judges actually employed.” Beebe, *supra* note 1, at 1607.

<sup>25</sup> *Id.*

<sup>26</sup> *Id.*

Table F

	Table F—All Cases Correlation		F 1 Strength of the marks					
	LoC found	LoC not found	LoC found	LoC not found	LoC outstanding issues	Supports LoC	Against LoC	Other
F1 Strength of the marks	LoC found	1						
	LoC not found	-0.8381*	1					
	LoC outstanding issues	-0.2593*	-0.3094*	1				
	Supports LoC	0.6081*	-0.5443*	-0.0963	1			
F2 Similarity of the marks	Against LoC	-0.4964*	0.5751*	-0.1530*	-0.7529*	1		
	Other	-0.2295*	0.0292	0.3483*	-0.4627*	-0.2351*	1	
	Supports LoC	0.7604*	-0.6536*	-0.1684*	0.4680*	-0.3983*	-0.1546*	1
	Against LoC	-0.6543*	0.7006*	-0.0998	-0.3586*	0.3896*	0.0047	-0.2873*
F3 Proximity of the products	Other	-0.2349*	-0.056	0.5087*	-0.2269*	0.0355	0.2873*	
	Supports LoC	0.5607*	-0.5157*	-0.0644	0.4238*	-0.3616*	-0.1389*	
	Against LoC	-0.4547*	0.5519*	-0.1844*	-0.3430*	0.3469*	0.0394	
	Other	-0.2173*	0.0283	0.3287*	-0.1651*	0.0713	0.1478*	
F4 Likelihood of bridging the gap	Supports LoC	0.3633*	-0.2684*	-0.1580*	0.2796*	-0.2478*	-0.0791	
	Against LoC	-0.4055*	0.4756*	-0.1351	-0.2044*	0.2573*	-0.0447	
	Other	0.0193	-0.1653*	0.2591*	-0.0777	0.0035	0.11	
	Supports LoC	0.5318*	-0.5077*	-0.0281	0.3546*	-0.2700*	-0.1601*	
F5 Evidence of actual confusion	Against LoC	-0.5694*	0.6121*	-0.0912	-0.3786*	0.3607*	0.0733	
	Other	0.1157	-0.1888*	0.1325	0.0759	-0.1406*	0.0773	
	Supports LoC	0.5195*	-0.4718*	-0.0702	0.3615*	-0.2997*	-0.1302	
	Against LoC	-0.4964*	0.6510*	-0.2873*	-0.3544*	0.4212*	-0.044	
F6 Defendant's bad faith/intent	Other	0.0444	-0.2843*	0.4260*	0.0417	-0.1900*	0.1943*	
	Supports LoC	0.2850*	-0.2369*	-0.0774	0.3315*	-0.2164*	-0.1981*	
	Against LoC	-0.2287*	0.2526*	-0.0486	-0.1559*	0.1981*	-0.0366	
	Other	-0.0404	-0.0195	0.1051	-0.1407*	0.01	0.1944*	
F7 Quality of the products	Supports LoC	0.3633*	-0.3581*	0.0007	0.3532*	-0.2370*	-0.2024*	
	Against LoC	-0.3926*	0.4998*	-0.2005*	-0.2226*	0.2711*	-0.0364	
	Other	0.0704	-0.1923*	0.2178*	-0.1051	-0.0621	0.2390*	
	Competing	0.3128*	-0.3206*	0.0224	0.2597*	-0.2304*	-0.0733	
F8 Consumer sophistication	Non-competing	-0.3158*	0.4129*	-0.1807*	-0.2442*	0.2812*	-0.0181	
	Other	-0.0218	-0.095	0.2063*	-0.0418	-0.0479	0.1262	
	Plaintiff's name	0.0479	-0.0772	0.0531	-0.0141	0.0486	-0.0445	
	Defendant's name	0.1069	-0.0599	-0.0803	0.0504	-0.0266	-0.0385	

n=206

Table F Continued

	F2 Similarity of the marks		F3 Proximity of the products		F4 Likelihood of bridging the gap		F5 Evidence of actual confusion		
	Supports LoC Against LoC	Other	Supports LoC Against LoC	Other	Supports LoC Against LoC	Other	Supports LoC Against LoC	Other	
1	1	1	1	1	1	1	1	1	
-0.8605*	-0.2188*	-0.1156	-0.7330*	-0.2364*	-0.1886*	-0.3587*	-0.5341*	-0.0094	-0.5733*
-0.3089*	-0.2852*	-0.1156	-0.4877*	-0.3288*	-0.1886*	-0.3587*	-0.5341*	-0.0094	-0.5733*
0.3384*	0.2402*	-0.0396	-0.7330*	-0.2364*	-0.1886*	-0.3587*	-0.5341*	-0.0094	-0.5733*
-0.2134*	0.0992	0.2159*	-0.4877*	-0.3288*	-0.1886*	-0.3587*	-0.5341*	-0.0094	-0.5733*
-0.2095*	-0.1846*	-0.1362	0.4274*	-0.3288*	-0.1886*	-0.3587*	-0.5341*	-0.0094	-0.5733*
0.2511*	0.2350	-0.0279	-0.6268*	0.6409*	0.0728	-0.5976*	-0.2087*	0.1288	-0.3508*
-0.2145*	-0.0346	0.1474*	0.1512*	-0.2527*	0.1083	-0.5976*	-0.2087*	0.1288	-0.3508*
-0.0432	-0.3060*	-0.1670*	0.2818*	-0.3051*	-0.0109	0.2084*	-0.2087*	0.1288	-0.3508*
0.3855*	0.3709*	0.1418*	-0.3690*	0.3545*	0.0721	-0.1796*	0.3393*	-0.11288	-0.5661*
-0.4355*	-0.116	0.006	0.1383*	-0.0983	-0.0714	-0.0044	-0.1778*	0.1567*	-0.5661*
0.11	-0.2901*	-0.1691*	0.3339*	-0.2599*	-0.1433*	0.2245*	-0.2649*	0.0242	0.3389*
0.3711*	0.4304*	-0.0296	-0.2860*	0.3745*	-0.0721	-0.0569	0.3406*	-0.2410*	-0.3545*
-0.4040*	-0.2115*	0.2186*	-0.0117	-0.1760*	0.2426*	-0.1730*	-0.1293	0.2677*	0.0664
0.092	-0.1871*	-0.1005	0.2786*	-0.2029*	-0.1376*	0.2200*	-0.1845*	-0.0408	0.2011*
0.2348*	0.2326*	0.0223	-0.1852*	0.1937*	0.0159	-0.019	0.1701*	-0.1289	-0.2502*
-0.2384*	-0.0436	0.0638	-0.0722	0.0026	0.0999	-0.1652*	0.0075	0.1432*	0.047
0.0091	-0.2367*	-0.0898	0.3256*	-0.2700*	-0.1187	0.2322*	-0.2489*	0.0035	0.1988*
0.2776*	0.3091*	-0.0558	-0.3318*	0.3937*	-0.0314	-0.1840*	0.4205*	-0.1945*	-0.2747*
-0.2722*	-0.1039	0.1493*	0.0411	-0.1635*	0.1510*	-0.028	-0.2134*	0.2086*	0.1038
0.0233	-0.1305	-0.067	0.6072*	-0.5092*	-0.2139*	0.0506	-0.5206*	0.4013*	0.1315
0.1622*	0.1756*	0.0019	-0.5047*	0.6144*	-0.0676	-0.0337	0.6139*	-0.4967*	-0.1602*
-0.1722*	-0.0486	0.0911	-0.1847*	-0.0965	0.3877*	-0.0264	-0.0799	0.0926	0.027
-0.0001	-0.0517	0.0728	-0.0168	-0.0294	0.0617	0.0423	0.0271	-0.0615	0.0583
0.0124	-0.0791	-0.0727	0.0071	0.0269	-0.0447	0.1005	0.0859	-0.1648*	0.0723
0.1151									0.0301
									-0.1071

Table G

Table G—PI/BT Correlation			F1 Strength of the marks				F2 Similarity of the marks			
	LoC found	LoC not found	LoC found	LoC not found	Supports LoC	Against LoC	Other	Supports LoC	Against LoC	Other
F1 Strength of the marks	LoC found	1								
	LoC not found	-1*	1							
	Supports LoC	0.6554*	-0.6554*	1						
F2 Similarity of the marks	Against LoC	-0.6112*	0.6112*	-0.8203*	1					
	Other	-0.1471	0.1471	-0.4038*	-0.1920*	1				
	Supports LoC	0.7683*	-0.7683*	0.5087*	-0.4658*	-0.1279	1			
F3 Proximity of the products	Against LoC	-0.7108*	0.7108*	-0.4203*	0.4221*	0.0461	-0.9251*	1		
	Other	-0.1890*	0.1890*	-0.2537*	0.1373	0.2157*	-0.2461*	-0.1403	1	
	Supports LoC	0.5474*	-0.5474*	0.4025*	-0.3999*	-0.051	0.3156*	-0.2805*	-0.1069	1
F4 Likelihood of bridging the gap	Against LoC	-0.4784*	0.4784*	-0.3470*	0.3290*	0.0693	-0.1538	0.1282	0.0738	
	Other	-0.1841*	0.1841*	-0.142	0.1634	-0.0178	-0.2787*	0.2597*	0.0637	
	Supports LoC	0.3349*	-0.3349*	0.3548*	-0.3110*	-0.1114	0.1854*	-0.1673*	-0.0564	
F5 Evidence of actual confusion	Against LoC	-0.4085*	0.4085*	-0.2637*	0.2562*	0.0255	-0.2002*	0.1707*	0.0865	
	Other	0.0194	-0.0194	-0.1277	0.0839	0.0849	-0.0107	0.018	-0.0179	
	Supports LoC	0.5420*	-0.5420*	0.3425*	-0.2831*	-0.1349	0.3913*	-0.3525*	-0.1206	
F6 Defendant's bad faith/intent	Against LoC	-0.6075*	0.6075*	-0.4335*	0.3384*	0.2025*	-0.4346*	0.3559*	0.2249*	
	Other	0.1244	-0.1244	0.1323	-0.0879	-0.0865	0.0855	-0.0384	-0.1249	
	Supports LoC	0.5282*	-0.5282*	0.3480*	-0.3487*	-0.0395	0.3652*	-0.3224*	-0.1293	
F7 Quality of the products	Against LoC	-0.5825*	0.5825*	-0.4356*	0.4101*	0.0915	-0.4609*	0.3978*	0.1863*	
	Other	0.1241	-0.1241	0.1503	-0.1158	-0.0727	0.1627	-0.1316	-0.0882	
	Supports LoC	0.2467*	-0.2467*	0.3290*	-0.2436*	-0.1749*	0.2055*	-0.1669*	-0.1099	
F8 Consumer sophistication	Against LoC	-0.1837*	0.1837*	-0.0132	0.034	-0.0317	-0.2166*	0.2229*	-0.0041	
	Other	-0.063	0.063	-0.2733*	0.1826*	0.1770*	-0.0006	-0.0378	0.0983	
	Supports LoC	0.4085*	-0.4085*	0.3287*	-0.2635*	-0.1427	0.3129*	-0.3039*	-0.0402	
F9 Products are competing	Against LoC	-0.4461*	0.4461*	-0.3009*	0.2640*	0.0941	-0.3027*	0.2835*	0.0657	
	Other	0.0778	-0.0778	-0.0019	-0.0238	0.0413	0.0163	-0.0042	-0.0318	
	Supports LoC	0.2677*	-0.2677*	0.2028*	-0.1758*	-0.0668	0.1434	-0.0972	-0.1258	
F10 Marks contain parties' name	Non-competing	0.2907*	-0.2907*	-0.1625	0.2074*	-0.0529	-0.1404	0.068	0.1926*	
	Other	0.0089	-0.0089	-0.0707	-0.0279	0.1659	-0.0158	0.0471	-0.0792	
	Plaintiff's name	0.1201	-0.1201	0.0669	-0.0346	-0.0594	0.0721	-0.059	-0.0373	
Defendant's name	0.0754	-0.0754	0.0882	-0.0539	-0.0653	0.0939	-0.0798	-0.041		

n=139

Table G Continued

	F3 Proximity of the products		F4 Likelihood of bridging the gap		F5 Evidence of actual confusion		F6 Defendant's bad faith/intent	
	Supports LoC	Against LoC	Other	Supports LoC	Against LoC	Other	Supports LoC	Against LoC
1								
-0.7906*	1							
-0.4682*	-0.1967*	1						
0.4787*	-0.3988*	-0.1887*	1					
-0.6970*	0.6526*	0.1731*	-0.3750*	1				
0.1213	-0.1605	0.0363	-0.6414*	-0.4708*	1			
0.2771*	-0.3108*	0.0038	0.1988*	-0.133	-0.0772	1		
-0.3400*	0.2734*	0.1479	-0.1831*	0.2217*	-0.0093	-0.5363*	1	
0.0955	0.0099	-0.164	0.0046	-0.1092	0.0861	-0.4039*	-0.5555*	1
0.3468*	-0.2637*	-0.1723*	0.2126*	-0.3133*	0.057	0.3096*	-0.3543*	-0.0391
-0.2938*	0.2675*	0.0836	-0.1018	0.2611*	-0.1192	0.3623*	-0.0391	-0.6938*
-0.0366	-0.0312	0.1001	-0.1253	0.0379	0.0879	-0.0457	-0.045	-0.3057*
0.2652*	-0.1995*	-0.1348	0.2301*	-0.1765*	-0.0729	0.2034*	-0.2178*	0.0356
-0.1195	0.088	0.0635	0.0073	0.161	-0.1402	0.1680*	0.0561	-0.1521
-0.1314	0.1004	0.0646	-0.2046*	0.021	0.1774*	0.0218	-0.0766	-0.2036*
0.3623*	-0.2967*	-0.164	0.2713*	-0.2626*	-0.0409	0.1579	-0.1405	0.3190*
-0.4517*	0.3671*	0.1910*	-0.2239*	0.4256*	-0.1392	0.2801*	-0.0032	0.3190*
0.1313	-0.1146	-0.0444	-0.029	-0.2041*	0.1965*	0.1199	-0.0541	-0.3004*
0.5690*	-0.4514*	-0.2562*	0.0401	-0.5167*	0.3895*	0.0807	-0.1675*	0.0086
-0.4669*	0.4700*	0.0694	-0.0502	0.6345*	-0.4773*	-0.0989	0.1694*	-0.0861
-0.1836*	0.0121	0.2715*	0.0103	-0.1154	0.0857	0.0177	-0.0618	0.0361
0.0564	-0.017	-0.0646	0.0253	-0.0074	-0.0179	0.2233*	-0.0877	-0.1249
0.0022	0.0479	-0.0711	-0.0022	0.0607	-0.0482	0.1037	-0.0392	-0.0598
								-0.0656
								-0.0785
								0.1840*

Table G Continued

	F7 Quality of the products Supports Loc Against Loc	F8 Consumer sophistication Supports Loc Against Loc	F9 Products are competing Other Competing Non-competing	F10 Marks contain parties' names Plaintiff's name Defendant's name
1	1	1	1	1
-0.2922*	-0.5639*	-0.2984*	-0.5637*	-0.5346*
-0.6250*	-0.0992	-0.2984*	-0.5637*	0.163
0.4394*	0.2675*	-0.0235	-0.3967*	0.157
-0.2256*	-0.1959*	0.3315*	-0.3967*	-0.7397*
-0.1988*	0.1724*	-0.15	-0.3930*	0.163
0.1157	0.128	-0.0046	0.3749*	-0.4312*
-0.0902	0.0417	0.0438	0.0566	-0.2883*
0.0698	-0.0992	0.0207	-0.0903	0.1105
0.0438	0.0652	-0.0911	0.0286	0.0309
			-0.0816	0.0138
			0.0763	0.5293*
			0.0138	1

*Stampeding*

*The Concept of Stampeding*

¶20 Beebe applies a “coherence-based reasoning model” to the decision-making process that courts undertake in determining whether a likelihood of confusion exists.<sup>27</sup> Judicial opinions, at least on their surface, evince a formalized multifactor analysis: a routine and methodical application of the factors, modified with familiar guiding principles—such as the exhortation that the factors are non-exhaustive<sup>28</sup> or that the test is not meant to be mechanically applied<sup>29</sup>—and synthesized into a final result, with each factor given appropriate weight. The coherence-based reasoning model hypothesizes a bi-directionality of decision-making, in which our cognitive processes cycle through mental models constructed from a network of variables.<sup>30</sup> Instead of logically proceeding in one direction from premises and facts to conclusions, conclusions affect and are affected in return by our premises and facts.<sup>31</sup> As we attempt to fit our mental model to the constraints imposed by these variables, our perceptions of the variables shift; these shifts in coherence determine and form the bases for our decisions.<sup>32</sup>

¶21 Applying this model to the multifactor likelihood of confusion test, Beebe hypothesizes that courts will tend to stampede factor outcomes to favor the test outcome when they find a likelihood of confusion, although this stampeding phenomenon will not be evident when courts do not find a likelihood of confusion.<sup>33</sup> This is partly because of the selection bias inherent in litigated cases; a plaintiff is unlikely to bring an action for trademark infringement unless the facts of the case are sufficiently favorable to win at least a few factors. His hypothesis is also grounded in the argument that a finding of no likelihood of confusion “represents an endorsement of the status quo, while a finding of a likelihood of confusion generally leads to an injunction intervention of the status quo.”<sup>34</sup> A finding of likelihood of confusion requires the appearance of a more decisive outcome in order to justify such an intervention.

¶22 Beebe also hypothesizes that the factors will affect each other, with the outcome of the “core factors” of intent and actual confusion stampeding the remainder of the test factors.<sup>35</sup> When intent and, to a lesser degree, actual confusion favor a likelihood of confusion, he theorizes that these outcomes will influence the judge’s perception of the other factors and increase the likelihood that he or she will conclude that the other factors favor a likelihood of confusion.<sup>36</sup>

¶23 In order to render his data in a form susceptible to empirical analysis of this stampeding phenomenon, Beebe calculated a “stampede score” for each opinion.<sup>37</sup> A Second Circuit opinion’s stampede score is the difference between the proportion of the eight factors that favored a finding of a likelihood of confusion and the proportion of the eight factors that did not favor a likelihood of confusion. In an opinion where all eight factors favor a likelihood of confusion, the stampede score is 1.000, whereas for an opinion in which all eight factors disfavor a likelihood of confusion, the stampede score is -1.000. A score of 0.000 represents an equal division of factors favoring and disfavoring a likelihood of confusion.

---

<sup>27</sup> Beebe, *supra* note 1, at 1615-22.

<sup>28</sup> *E.g.*, *Physicians Formula Cosmetics, Inc. v. West Cabot Cosmetics, Inc.*, 857 F.2d 80, 85 (2d Cir. 1988).

<sup>29</sup> *E.g.*, *N.Y. Stock Exch., Inc. v. N.Y., N.Y. Hotel, LLC*, 293 F.3d 550, 555 (2d Cir. 2002) (citing *Nora Bevs. Inc. v. The Perrier Group of Am., Inc.*, 269 F.3d 114, 119 (2d Cir. 2001)).

<sup>30</sup> *See* Beebe, *supra* note 1, at 1615-17 (summarizing this model).

<sup>31</sup> *Id.*

<sup>32</sup> *Id.*

<sup>33</sup> *Id.* at 1614-15.

<sup>34</sup> *Id.* at 1619.

<sup>35</sup> *Id.* at 1620.

<sup>36</sup> *Id.* at 1620-21.

<sup>37</sup> *Id.* at 1618.

*Results of the Stamped Analysis*

¶24 Beebe's national data supported his hypothesis that judges routinely stamped factor outcomes to favor the test outcome, particularly when they found a likelihood of confusion.<sup>38</sup> Our data tend not to support this hypothesis, which is consistent with Beebe's own finding that Second Circuit district courts were far less prone to stampede the factors than the other courts he sampled.<sup>39</sup> While Beebe found a mean stampede score of .723<sup>40</sup> in national preliminary injunction and bench trial opinions favoring a likelihood of confusion, we observed a lower mean stampede score of .570<sup>41</sup> in S.D.N.Y. preliminary injunction and bench trial opinions favoring a likelihood of confusion. Our mean stampede score, however, was very similar to that observed by Beebe in his Second Circuit district court opinions, .593.<sup>42</sup>

¶25 Inconsistent with Beebe's hypothesis, we did not observe a greater stampeding effect in opinions finding a likelihood of confusion. Table H sets forth the mean stampede scores for our sample of 206 opinions, sampled by outcome and posture. We observed that our mean stampede score was fairly consistent against posture. Even on cross summary judgment motions, opinions that found a likelihood of confusion yielded a stampede score of 0.604.<sup>43</sup> Figure 4 reports the distribution of stampede scores by disposition for our 139 preliminary injunction and bench trial opinions. We observed fewer instances of stampede scores of 1.000 or -1.000 than did Beebe, which is also inconsistent with his hypothesis. Of the opinions favoring a likelihood of confusion, 10.6% yielded a stampede score of 1.000, whereas of the opinions disfavoring a likelihood of confusion, 3.9% yielded a stampede score of -1.000, which somewhat tends to support his hypothesis.

¶26 Consistent with Beebe's results, we observed a higher occurrence of irregular stampede scores in opinions that disfavored a likelihood of confusion than those that favored a likelihood of confusion. An irregular stampede score for an opinion that found no likelihood of confusion represents an opinion where a majority of factors favored a likelihood of confusion. This observation is consistent with the guidance that courts should not apply the *Polaroid* test mechanically.

¶27 Our data are also less supportive of Beebe's second hypothesis that when intent and, to a lesser degree, actual confusion favor a likelihood of confusion, we should observe greater stampeding of the rest of the test factors. In addition to plaintiff win rates, Table B reports the mean stampede score by factor outcome in our set of preliminary injunction and bench trial opinions. While the highest mean stampede scores are observed when intent<sup>44</sup> and, to a lesser degree, actual confusion<sup>45</sup> favor a likelihood of confusion, these stampede scores are neither as high as Beebe's, nor particularly high as compared to stampede scores yielded under other factor outcomes. However, we do observe, consistent with Beebe's results, severe negative skewing<sup>46</sup> and peakedness,<sup>47</sup> characteristics that are unique to these factors.

¶28 Interestingly, the highest stampede scores can be observed by varying both factor and test outcome. Table C reports the mean stampede scores by test outcome in our set of preliminary injunction and bench trial opinions. As opposed to Beebe's core factors of intent and actual confusion, we find that, in preliminary injunction and bench trial opinions that favored a likelihood of confusion, the highest stampede scores are yielded by opinions where the quality of the products,<sup>48</sup> and, to a lesser degree, the consumer sophistication

---

<sup>38</sup> *Id.* at 1618-19.

<sup>39</sup> *Id.* at 1621.

<sup>40</sup> SD=.291, n=102.

<sup>41</sup> SD=.267, n=68.

<sup>42</sup> SD=.315, n=26.

<sup>43</sup> SD=.243, n=6.

<sup>44</sup> SS=0.613, SD=0.340, n=43.

<sup>45</sup> SS=0.587, SD=0.376, n=39.

<sup>46</sup> Skewing here refers to the shape of the distribution of stampede scores. A negatively skewed distribution is one in which low values are rare. For example, plotting the grades from an easy math test on a graph will result in a negatively skewed distribution. *See* W. PAUL VOGT, *DICTIONARY OF STATISTICS AND METHODOLOGY: A NONTECHNICAL GUIDE FOR THE SOCIAL SCIENCES* 297 (3d ed. 2005).

<sup>47</sup> Peakedness describes the *kurtosis* of the distribution. Opinions finding that the intent or actual confusion factors favored a likelihood of confusion were characterized by stampede scores with comparatively extreme kurtosis. A high kurtosis score here means that the distribution of stampede scores was extremely pointy, as compared to a bell-shaped or normal curve. Negative kurtosis scores, as observed in nearly every other stampede score distribution set out in the table, represent a flatter curve. *Id.*

<sup>48</sup> SS=0.807, SD=0.173, n=24.

factor,<sup>49</sup> favored a likelihood of confusion. This observation is not true of Beebe’s data, and is unexpected given that neither factor is strongly correlated with test outcome or appears on our classification trees. Possibly, these factors are not independently important enough to affect the outcome of the likelihood of confusion test. Indeed, they are *so* irrelevant to the analysis that when courts decide these factors in favor of the plaintiff, they are likely stampeding the factors in favor of a likelihood of confusion rather than substantively analyzing the factors. This possibility is somewhat supported by the observation that the quality of the products factor was often not explicitly decided in favor of any result.<sup>50</sup>

Table H

MULTIFACTOR TEST STAMPEDE SCORES BY DISPOSITION AND POSTURE IN 206 DISTRICT COURT OPINIONS (SS = stampede score)									
Disposition	Posture	N	Mean SS	SD	Min. SS	Median SS	Max. SS	N With [SS] of 1.000	N with Irregular SS*
LoC Found	Preliminary Injunction†	52	0.565	0.241	0.000	0.563	1.000	4	0
	Bench Trial	17	0.581	0.336	-0.125	0.625	1.000	3	1
	Plaintiff SJ Motion	8	0.484	0.245	0.250	0.438	0.875	0	0
	Defendant SJ Motion	2	0.625	0.530	0.250	0.625	1.000	1	0
	Cross SJ Motions	6	0.604	0.243	0.375	0.563	1.000	1	0
LoC Not Found	Preliminary Injunction‡	41	-0.357	0.299	-0.875	-0.375	0.500	0	3
	Bench Trial	31	-0.403	0.319	-1.000	-0.375	0.250	1	4
	Plaintiff SJ Motion	2	-0.438	0.265	-0.625	-0.438	-0.250	0	0
	Defendant SJ Motion	22	-0.608	0.257	-1.000	-0.625	-0.125	3	0
	Cross SJ Motions	7	-0.446	0.374	-0.875	-0.500	0.250	0	1
SJ Denied	Plaintiff SJ Denied	4	0.000	0.102	-0.125	0.000	0.125	0	---
	Defendant SJ Denied	3	0.125	0.500	-0.375	0.125	0.625	0	---
	Cross SJs Denied	11	0.182	0.258	-0.250	0.250	0.500	0	---
Total		206	0.029	0.556	-1.000	0.000	1.000	13	---

†Includes one TRO, *Energybrands, Inc. v. Beverage Marketing USA, Inc.*, 2002 WL 826814

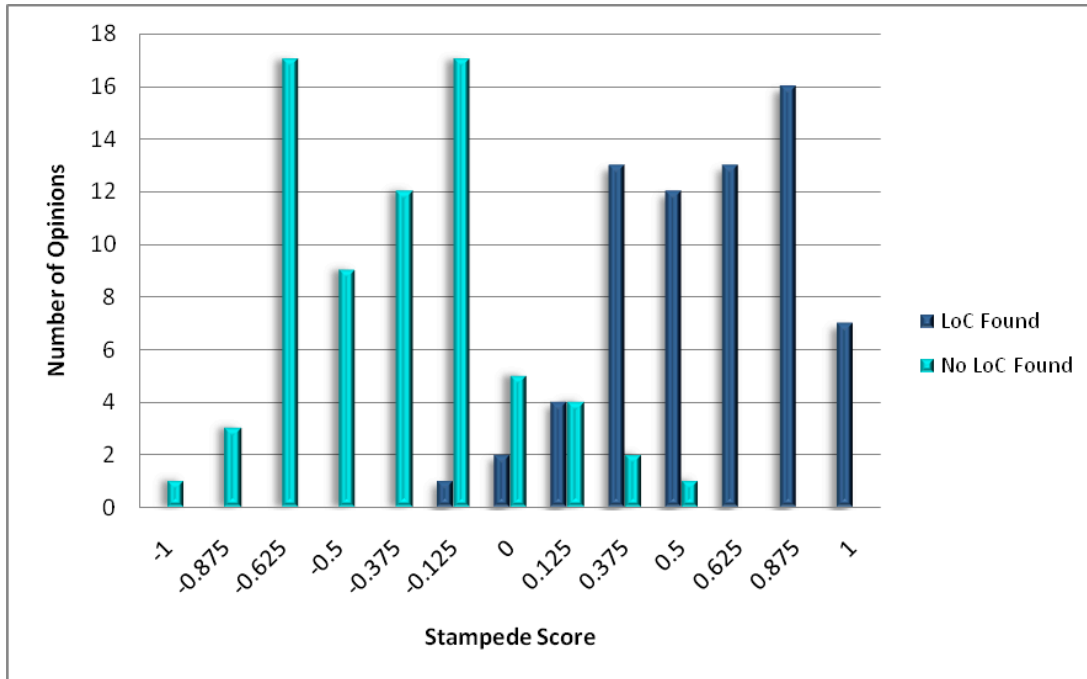
‡Includes one Magistrate Report, *V&S Vin & Spirit Aktiebolag (publ) v. Absolute Publishing USA Inc.*, 2005 WL 3272828

\*"Irregular Count" describes opinions in which the winner of the multifactor test won fewer factors than the loser of the test.

<sup>49</sup> SS=0.708, SD=0.225, n=33.

<sup>50</sup> In our set of 139 bench trial and preliminary injunction opinions, seventy-six opinions did not explicitly decide the quality of the products factor in favor of one outcome (54.7%). The only factor to come close in infrequency of outcome determination was the bridging the gap factor (44.6%), which was often overlooked or found to be neutral because the products or services at issue were in direct competition.

Figure IV: The Distribution of Multifactor Stampede Scores in 139 Preliminary Injunction and Bench Trial Opinions



*Historical Trends*

¶29 The cases in each of the all cases and PI/BT datasets were grouped into three-year baskets, and relevant metrics assessed for each of the time periods in order to analyze historical trends.<sup>51</sup> The major caveat in these historical trends data is that, due to the relatively small number of cases analyzed, statistical significance of these findings cannot be guaranteed. However, to the extent that these trends may reflect actual trends in the data, and not simply random variations, several interesting observations are revealed.

¶30 A look at the historical trends table reveals a general lack of significant trends, except possibly in three areas. First, at least in the all cases dataset, it appears that the rate at which courts find a likelihood of confusion has declined. In our first time period, from 1993 to 1996, the ratio of finding a likelihood of confusion to no likelihood of confusion was about 1:1. In our last two time periods, the corresponding ratio has decreased to between a little less than 1:2 and 2:3. This difference has not been accompanied by an increase in the number of cases in which the court found outstanding issues of fact; such cases have held steady at three or four per time period.

¶31 Second, the data shows a trend away from subscribing to any particular theory—either similarity, tarnishment, or both—when analyzing the quality factor. While the tarnishment theory has dominated this factor in all time periods, the number of cases subscribing to the theory has declined. In the first period’s all cases dataset, courts that articulated a particular theory for the quality factor subscribed to tarnishment more than 2:1, compared to similarity or a combination of similarity and tarnishment. In the final period, however, while tarnishment is still the most-mentioned theory, the ratio is about 4:3. The corresponding results in our PI/BT dataset are even more pronounced, with the ratios at 2:1 in the first period falling to about 1:1 in the last period.

¶32 Finally, there appears to be a trend in the relationship between inherent or commercial strength of the mark (sub-factors under the strength of the mark factor) and the ultimate finding of likelihood of confusion. For example, in the first period of the all cases dataset, two-thirds of the cases finding inherent strength also found a likelihood of confusion. This percentage drops to about half in the final period. The numbers are similar for commercial strength.

*Factor 1: The Strength of the Mark*

¶33 The first factor that courts consider is the strength of the mark.<sup>52</sup> The analysis reveals this factor’s importance in deciding likelihood of confusion. Strength of the mark was the second most useful inquiry in one- and two-node decision trees.<sup>53</sup> This is supported by multiple statistically significant correlations. There is a positive correlation between a finding that the strength of the mark favors confusion and an ultimate finding of likelihood of confusion.<sup>54</sup> This correlation was slightly more pronounced when opinions were limited to those ruling on a preliminary injunction or bench trial.<sup>55</sup> The inverse of this statement was also true: a finding that the strength of the mark does not favor confusion positively correlated with an ultimate finding of no likelihood of confusion.<sup>56</sup>

¶34 In the Second Circuit, the strength of the mark is separated into two categories: inherent strength (the distinctiveness of the mark itself) and commercial strength (the “secondary meaning” the mark has acquired in the marketplace).<sup>57</sup> Beebe found that, while the two inquiries are independent, commercial strength trumps inherent strength when the two metrics favor opposite holdings.<sup>58</sup> This conclusion was corroborated by our analysis.

---

<sup>51</sup> These tables are included as Table 6 in the Appendix.

<sup>52</sup> *Star Indus. v. Bacardi & Co.*, 412 F.3d 373, 384 (2d Cir. 2005).

<sup>53</sup> *See supra* p. 14.

<sup>54</sup> Coefficient of 0.6081.

<sup>55</sup> Coefficient of 0.6554.

<sup>56</sup> Coefficient of 0.5751 for all cases, 0.6112 for preliminary injunction/bench trial cases.

<sup>57</sup> *Nat’l Distillers Prods. Co., LLC v. Refreshment Brands, Inc.*, 198 F. Supp. 2d 474, 480 (S.D.N.Y. 2002).

<sup>58</sup> Beebe, *supra* note 1, at 1636.

*Inherent Strength*

¶35 Of the 206 total cases, 122 found a mark to be inherently distinctive (59.22%), forty-two found a mark to be inherently weak (20.39%) and forty-two were neutral or unclear (20.39%). These findings were important to both the determination of the first factor and the determination of confusion in general. There is a statistically significant positive correlation between a finding that the mark is inherently strong and a finding that the strength of the mark favors confusion.<sup>59</sup> Interestingly, there were some statistically significant correlations between a finding of inherent strength and finding that *other Polaroid* factors favor a likelihood of confusion. Specifically, there was (1) a negative correlation between a finding of inherent weakness and a finding that the similarity of the marks favors confusion,<sup>60</sup> and (2) a positive correlation between a finding of inherent strength and a finding that the proximity of the goods favors confusion.<sup>61</sup> Additionally, the proportion of factor outcomes favoring likelihood of confusion was higher for inherently distinctive marks than for all marks examined.<sup>62</sup> Thus, a significant positive correlation between a finding of inherent strength and an ultimate finding of a likelihood of confusion is unsurprising.<sup>63</sup>

*Abercrombie Factors*

¶36 Inherent strength is determined in the Second Circuit by placing a mark in one of five categories: fanciful, arbitrary, suggestive, descriptive, or generic.<sup>64</sup> While the *Abercrombie* case first announced these categories, it listed "arbitrary or fanciful" as one category.<sup>65</sup> Later cases have separated the two into different categories, though some decisions still find a mark to be "arbitrary or fanciful."<sup>66</sup>

¶37 Fanciful marks have no independent meaning and are usually invented specifically for use as a mark.<sup>67</sup> Arbitrary marks have independent meaning, but that meaning is unrelated to their use as marks.<sup>68</sup> Suggestive marks "do not describe but merely suggest the features of the product, requiring the purchaser to use imagination, thought and perception to reach a conclusion as to [its] nature."<sup>69</sup> Descriptive marks describe the products to which the mark is attached.<sup>70</sup> Finally, generic marks are either synonymous with the "genus of which the particular product is a species," or are "common description[s]" of the product.<sup>71</sup>

¶38 Arbitrary, fanciful, and suggestive marks are considered inherently distinctive and are afforded the most stringent trademark protection.<sup>72</sup> Descriptive marks are considered inherently weak and are only protected if they have acquired "secondary meaning" (*i.e.*, they have commercial strength).<sup>73</sup> Generic marks are inherently weak and never receive trademark protection.<sup>74</sup> Beebe found that district courts make little use of the *Abercrombie* inquiry, with only 164 of 331 cases (49.55%) placing the mark into one of the five categories.<sup>75</sup> This conclusion is supported by the fifteen-year S.D.N.Y. sample as well, with only 103 of 206 cases (50%) placing the mark into one of the five categories.<sup>76</sup> Of those 103 cases, three (2.91%) labeled marks as

---

<sup>59</sup> Coefficient of 1.377791, z score of 4.57,  $P > |z| = 0.000$ .

<sup>60</sup> Coefficient of -1.090.

<sup>61</sup> Coefficient of 0.740.

<sup>62</sup> See Appendix, Table 7.

<sup>63</sup> Coefficient of 4.90.

<sup>64</sup> *Abercrombie & Fitch Co. v. Hunting World, Inc.*, 537 F.2d 4, 9 (2d Cir. 1976).

<sup>65</sup> *Id.*

<sup>66</sup> See *infra* note 76.

<sup>67</sup> *Solow v. BMW (US) Holding Corp.*, No. 97 Civ. 1373, 1998 U.S. Dist. LEXIS 16059, at \*10 (S.D.N.Y. Oct. 14, 1998).

<sup>68</sup> *Id.*

<sup>69</sup> *Id.*, quoting *W.W.W. Pharm. Co. v. Gillette Co.*, 984 F.2d 567, 572 (2d Cir. 1993).

<sup>70</sup> *Id.*

<sup>71</sup> *Id.*

<sup>72</sup> *Id.* at \*10-11.

<sup>73</sup> *Id.* at \*11.

<sup>74</sup> *Id.*

<sup>75</sup> Beebe, *supra* note 1, at 1635-36.

<sup>76</sup> In order to have placed a mark into one of the categories, a court must have explicitly given the mark one category and one category only. The not uncommon practice of labeling a mark as "fanciful or arbitrary" was coded as "other" in our analysis.

“fanciful,” twenty-three (22.33%) as “arbitrary,” forty-eight (44.60%) as “suggestive,” twenty-seven (26.21%) as “descriptive,” and two (1.94%) as “generic.”

¶39 Beebe’s conclusion about lack of use was also bolstered by the fact that, while all of the *Abercrombie* factors either perfectly predicted or significantly correlated with inherent strength, none of the factors correlated significantly with a finding that the first factor favors a likelihood of confusion.

¶40 Interestingly, there were significant correlations between some of the *Abercrombie* categories and an ultimate finding of a likelihood of confusion: arbitrary marks were positively correlated with an ultimate finding of a likelihood of confusion<sup>77</sup> and descriptive marks were negatively correlated with an ultimate finding of a likelihood of confusion.<sup>78</sup>

¶41 This may be explained by the fact that there were some significant correlations between certain *Abercrombie* categories and other *Polaroid* factors. Finding that a mark is suggestive or descriptive negatively correlated with a finding that the similarity of the marks favors confusion.<sup>79</sup> This may be because, since suggestive and descriptive marks are entitled to lesser protection than arbitrary or fanciful marks, courts require a higher degree of similarity with infringing marks in order to find that that factor favors a likelihood of confusion. A finding that a mark is descriptive was also negatively correlated with a finding that there is evidence of actual confusion.<sup>80</sup>

¶42 Like the correlation with similarity of the marks, this correlation may imply that courts require more proof of actual confusion when the marks are descriptive than when they are inherently strong. Additionally, the proportion of factor outcomes favoring a likelihood of confusion was higher than the average plaintiff win rate for fanciful and arbitrary marks, while the proportion of factor outcomes favoring a likelihood of confusion was lower than the average plaintiff win rates for descriptive and generic marks.<sup>81</sup>

#### *Commercial Strength*

¶43 Of the 206 total cases, ninety-one (44.17%) were found to be commercially strong, fifty-three (25.73%) were found to be commercially weak, and sixty-two (30.10%) were either commercially neutral or unclear. The importance of commercial strength to all aspects of the likelihood of confusion inquiry is striking, especially when compared to inherent strength. A finding of commercial strength is positively correlated with a finding that the first factor favors a likelihood of confusion.<sup>82</sup> Additionally, a finding of commercial strength is positively correlated with an ultimate finding of a likelihood of confusion.<sup>83</sup> Finally, commercial strength significantly affects the other *Polaroid* factors. A finding of commercial strength is positively correlated with (1) a finding that the similarity of the marks favors a likelihood of confusion<sup>84</sup> and (2) a finding that the quality of the goods favors a likelihood of confusion.<sup>85</sup> More importantly, a finding of commercial weakness is damning for a mark, as evidenced by its negative correlation with findings that (1) the similarity of the marks,<sup>86</sup> (2) the proximity of the goods,<sup>87</sup> (3) evidence of actual confusion,<sup>88</sup> (4) evidence of bad faith,<sup>89</sup> and (5) the sophistication of the consumers<sup>90</sup> favor a likelihood of confusion. This is supported by the fact that, when a mark is commercially weak, the plaintiff win rate and proportion of factor outcomes favoring a

---

<sup>77</sup> Coefficient of 1.062.

<sup>78</sup> Coefficient of -1.017.

<sup>79</sup> Coefficients of -1.448 for suggestive, -1.355 for descriptive.

<sup>80</sup> Coefficient of -1.662.

<sup>81</sup> See Appendix, Table 9.

<sup>82</sup> Coefficient of 3.865543.

<sup>83</sup> Coefficient of 6.14.

<sup>84</sup> Coefficient of 1.019.

<sup>85</sup> Coefficient of 1.112.

<sup>86</sup> Coefficient of -1.188.

<sup>87</sup> Coefficient of -1.366.

<sup>88</sup> Coefficient of -1.363.

<sup>89</sup> Coefficient of -1.449.

<sup>90</sup> Coefficient of -1.085.

likelihood of confusion are much lower than the relevant averages. By contrast, when a mark is commercially strong, the win rate and factor outcomes favoring a likelihood of confusion are all higher than average.<sup>91</sup>

*When Inherent and Commercial Strength Point in Different Directions*

¶44 Beebe found that, when inherent and commercial strength pointed in different directions, “a finding of acquired strength (or weakness) almost invariably trumped a finding of inherent weakness (or strength).”<sup>92</sup> This conclusion is supported by the fifteen-year S.D.N.Y. data set. When there is commercial strength, but no inherent strength (that is, inherent weakness or neutral or unclear inherent strength), the first factor favored a likelihood of confusion 87.10% of the time (twenty-seven instances), did not favor a likelihood of confusion 6.45% of the time (two instances), and was neutral or unclear 6.45% of the time (two instances). By contrast, when there is inherent strength, but no commercial strength (that is, commercial weakness or neutral or unclear commercial strength), the first factor favored confusion 46.77% of the time (twenty-nine instances), did not favor confusion 38.71% of the time (twenty-four instances), and was neutral or unclear 14.52% of the time (nine instances). This strong preference for commercial strength over inherent strength was true regardless of which, if any, *Abercrombie* category the court gave to the mark.

¶45 Our analysis of courts’ treatment of the first *Polaroid* factor supports Beebe’s conclusions. Inherent strength is important, but not nearly as important as commercial strength. *Abercrombie* factors are only fully used half of the time and, when they are used, they do not carry much weight. The evidence also supports the general proposition that “firms should choose inherently distinctive marks.”<sup>93</sup> Finally, the evidence supports Beebe’s suggestion that, by discussing inherent strength with the same importance as commercial strength in trademark opinions, judges are distorting the fact that commercial strength is more important to an ultimate finding of a likelihood of confusion.<sup>94</sup>

*Factor 2: The Similarity of the Marks*

¶46 The second factor that courts consider is the similarity between the two marks.<sup>95</sup> “In assessing similarity, courts look to the overall impression created by the logos and the context in which they are found, and then consider the totality of factors that could cause confusion among prospective purchasers.”<sup>96</sup> Beebe found similarity to be “by far the most important factor in the multifactor test.”<sup>97</sup> Our research supports his conclusion. Of the 113 cases in which the similarity of the marks favored a likelihood of confusion, eighty-five ultimately found a likelihood of confusion (75.22%). More tellingly, none of the seventy-eight cases in which the similarity of the marks did not favor a likelihood of confusion found an ultimate likelihood of confusion. A finding that the similarity of the marks favors a likelihood of confusion was significantly and positively correlated with an ultimate finding of confusion for all cases, as well as for the preliminary injunction/bench trial subset.<sup>98</sup> Beebe’s conclusion was borne out in the classification tree analysis as well, as it was the most useful factor for both one-node, two-node, and five-node trees.<sup>99</sup>

¶47 The importance of this factor has contrasting implications. On one hand, it seems to be a sign that the *Polaroid* test is acting as it should. Courts must ensure that trademark law is not used as a weapon against legitimate competition. By putting heavy weight on whether the competing marks are similar, courts draw a sharp line between the types of competition that are to be encouraged and the types that are off-limits. On the other hand, the importance of this factor also suggests that, when courts say that “[n]o single factor in the list is dispositive,”<sup>100</sup> such an assertion is not reflected in their analyses. What they mean to say is “no single

<sup>91</sup> See Appendix, Table 8.

<sup>92</sup> Beebe, *supra* note 1 at 1636.

<sup>93</sup> *Id.* at 1638.

<sup>94</sup> *Id.* at 1639.

<sup>95</sup> See *Star Indus. v. Bacardi & Co.*, 412 F.3d 373, 384 (2d Cir. 2005).

<sup>96</sup> *Id.* at 386 (citing *Gruner & Jahr USA Publ’g v. Meredith Corp.*, 991 F.2d 1072, 1078 (2d Cir. 1993)).

<sup>97</sup> Beebe, *supra* note 1, at 1623.

<sup>98</sup> Coefficients of 0.7006 (all cases) and 0.7108 (PT/BI cases).

<sup>99</sup> See *supra* pp. 15-16.

<sup>100</sup> *Regal Jewelry Co. v. Kingsbridge Int’l*, 999 F. Supp. 477, 491 (S.D.N.Y. 1998).

factor in the list is dispositive *so long as there is a threshold amount of similarity between the marks.*” A re-working of the likelihood of confusion test that first asks “is there a threshold amount of similarity between the two marks?” and then, if the answer is affirmative, moves on to the other seven *Polaroid* factors would comport with reality more closely than the current formulation.

*Factor 3: The Proximity of the Products*

¶48 The typical formulation of the proximity factor is as follows: “The third factor addresses whether, due to the commercial proximity of the competitive products, consumers may be confused as to their source. The key here is the possibility of confusion in the context in which consumers encounter, and consider purchasing, the parties’ products.”<sup>101</sup>

¶49 Our analysis shows a moderate correlation between the outcome of this factor’s analysis and an ultimate finding of a likelihood of confusion. In our PI/BT cases dataset, a finding that the proximity of the products factor favors a likelihood of confusion correlates with an ultimate finding of a likelihood of confusion with a value of 0.5474. Conversely, a finding that this factor militates against a likelihood of confusion correlates with an ultimate finding of no likelihood of confusion with a value of 0.4784. In our all cases dataset, a finding that the proximity of the products factor favors a likelihood of confusion correlates with an ultimate finding of a likelihood of confusion with a value of 0.5607. Conversely, a finding that this factor militates against a likelihood of confusion correlates with an ultimate finding of no likelihood of confusion with a value of 0.5519.

¶50 In our classification trees, proximity is a relatively important factor, useful in the best two-node classification tree, and is one of the five factors that is useful in the five-node classification tree. This factor is most useful if proximity disfavors confusion—that is, the data reveals that courts rely heavily on this factor *disfavoring* a likelihood of confusion when making an ultimate finding that there is no likelihood of confusion.

¶51 The relationship between the proximity factor favoring a likelihood of confusion and the question of whether the products are competing is interesting. The two inquiries should intuitively be similar, since they both consider the degree to which the products in question would be seen or sold together, thereby confusing consumers. In the dataset, where the products were found to be competing, it was also almost always found that this proximity factor favors a likelihood of confusion. Conversely, if a product is non-competing, there is still a reasonable (25%) chance it will be found that proximity favors a likelihood of confusion. Approaching the inquiry from the opposite angle, if proximity favors a likelihood of confusion, there is a good chance the products will be competing. If, however, the proximity factor militates against a likelihood of confusion, it is even more likely that the products will not be found to be competing. These findings confirm that courts view the proximity of the products factor as being similar to the question of whether the products are competing.

*Factor 4: Bridging the Gap*

¶52 The typical formulation of the bridging the gap factor is as follows:

This *Polaroid* factor examines the likelihood that the senior user will enter the junior user's market and compete with the junior user (“bridge the gap”). The factor weighs in the senior user’s favor either if bridging the gap is actually probable or if an average consumer perceives it as probable. The actual probability of bridging the gap is relevant because trademark law protects, in part, the senior user's interest in being able to expand into a related field in the future. The perception of consumers, on the other hand, affects the likelihood of confusion, because if consumers perceive bridging the gap as probable, they are more likely to believe that the junior user's products emanate from the senior user.<sup>102</sup>

---

<sup>101</sup> 24 Hour Fitness USA, Inc. v. 24/7 Tribeca Fitness, LLC, 447 F. Supp. 2d 266, 275-76 (S.D.N.Y. 2006) (citations omitted).

<sup>102</sup> Paco Sport, Ltd. v. Paco Rabanne Parfums, 86 F. Supp. 2d 305, 318 (S.D.N.Y. 2000) (citations omitted).

¶53 Our analysis found a moderate correlation between the outcome of this factor and an ultimate finding of a likelihood of confusion. Specifically, in the PI/BT cases dataset alone, a finding that the bridging the gap factor favors a likelihood of confusion correlates with an ultimate finding of a likelihood of confusion with a value of 0.3349. Conversely, a finding that this factor militates against a likelihood of confusion correlates with an ultimate finding of no likelihood of confusion with a value of 0.4085. Likewise, in the all cases dataset, a finding that the bridging the gap factor favors a likelihood of confusion correlates with an ultimate finding of a likelihood of confusion with a value of 0.3633. Conversely, a finding that this factor militates against a likelihood of confusion correlates with an ultimate finding of no likelihood of confusion with a value of 0.4756.

¶54 Courts generally analyze this factor in a straightforward manner. When the senior user is likely to enter the same market as the junior user — *i.e.*, bridge the gap between the two markets — such a finding supports a likelihood of confusion on this factor. Conversely, when courts find that the senior user will probably not enter the junior user’s market, such a finding militates against a likelihood of confusion. An interesting observation arises, however, in those cases involving already directly competing products or services—when there is no gap to bridge. In such cases, courts were somewhat divided as to whether the lack of a gap to bridge favors a likelihood of confusion or renders the factor irrelevant to the inquiry.<sup>103</sup>

¶55 In our all cases dataset, a modest majority of cases (fifty-two of eighty-one) treated a lack of a gap to bridge as rendering the entire bridging the gap factor neutral or irrelevant, while less than half of such cases (twenty-nine of eighty-one) treated the situation as one favoring a likelihood of confusion on this factor. In our “PI/BT” dataset, a slight majority (thirty-seven of fifty-nine) treated a no-gap finding as rendering this factor as neutral or irrelevant, while a minority (twenty-two of fifty-nine) treated the situation as favoring a likelihood of confusion. Furthermore, in that same dataset, courts that treat no gap to bridge as favoring a likelihood of confusion had a higher rate of finding a likelihood of confusion and a higher stampede score than those that treated no gap to bridge as neutral or irrelevant. No courts treated a finding of a lack of a gap to bridge as militating against confusion.

¶56 Another interesting observation is that some courts tend to lump this factor in with the proximity of the products factor such that the courts title their analysis of these two factors along the lines of “Proximity of the Products and Bridging the Gap.” In cases like these, a typical discussion of the two combined factors begins with: “The factor addresses ‘whether due to the commercial proximity of the competitive products, consumers may be confused as to their source.’” Where the parties compete in the same market, there is “no gap to bridge” and that factor tips in favor of a likelihood of confusion.<sup>104</sup>

¶57 This suggests that courts may view these two factors as going hand in hand. However, from our analysis of PI/BT cases, a finding of a situation favoring confusion in one of these two factors correlates with the same finding in the other factor only moderately, with a value of 0.4787. Conversely, a finding of militating against confusion in one of these two factors correlates with the same finding in the other factor somewhat more strongly, with a value of 0.6526. Likewise, from our all cases dataset, a finding of a situation favoring confusion in one of these two factors correlates with the same finding in the other factor only moderately, with a value of 0.4274. Conversely, a finding of militating against confusion in one of these two factors correlates with the same finding in the other factor somewhat more strongly, with a value of 0.6409.

¶58 The correlation between these two factors when they both favor confusion is somewhat muddled by the inconsistent way in which courts treat a finding of no gap to bridge (as mentioned above). This probably is why we see a stronger correlation between these two factors when they both militate against confusion, where the problem of courts’ inconsistent treatment of no-gap-to-bridge cases does not exist. Nevertheless, even in the latter case, there is only a moderate correlation, which somewhat refutes the view that these two factors go hand in hand.

---

<sup>103</sup> This ambiguity likely stems from a lack of direction as to how to apply this factor in cases involving competing products or services. The *Polaroid* test was originally applied exclusively to cases involving non-competing goods or services, and only later became the Second Circuit’s single likelihood of confusion test. *See* Beebe, *supra* note 1, at 1588 n. 33.

<sup>104</sup> *Artisan Mfg. Corp. v. All Granite & Marble Corp.*, 559 F. Supp. 2d 442, 451 (S.D.N.Y. 2008) (citations omitted).

*Factor 5: Actual Confusion*

¶59 “For purposes of the Lanham Act, actual confusion means ‘consumer confusion that enables a seller to pass off his goods as the goods of another.’”<sup>105</sup> Our findings on the relationship between the outcome of the actual confusion factor and the outcome of the a likelihood of confusion analysis were consistent with Beebe’s observations. In our five-node classification tree, a finding that the actual confusion factor favored a likelihood of confusion was useful in sorting cases. In opinions where the similarity factor favored confusion and the intent factor did not, a finding that the actual confusion factor favored confusion was sufficient to result in a finding of likelihood of confusion.

¶60 Of the preliminary injunction and bench trial opinions only, a finding that the actual confusion factor favored a likelihood of confusion correlated with an ultimate finding of likelihood of confusion with a value of 0.5420. Conversely, a finding that this factor militated against a likelihood of confusion correlates with an ultimate finding of no likelihood of confusion with a value of 0.6075.

¶61 Only three of the thirty-nine opinions in which the court found that the actual confusion factor favored a likelihood of confusion resulted in a finding of no likelihood of confusion. In all three, both the strength and similarity of the mark factors weighed against the finding of a likelihood of confusion.

¶62 The oft-cited principle that no showing of actual confusion is necessary to establish a likelihood of confusion<sup>106</sup> was arguably inconsistent with our data, with only eight of fifty-nine opinions in which the court found that this factor did not favor a finding of a likelihood of confusion nevertheless resulting in an overall finding of a likelihood of confusion, a plaintiff win rate of 13.6%.

¶63 While opinions in which the plaintiff cannot establish actual confusion are perhaps more likely to involve weak claims, given the recognized difficulty of establishing actual confusion,<sup>107</sup> this disparity in win rates suggests two possibilities, each tending to undermine the routine method of inquiry suggested by the explicitly formal analysis of the multifactor test as it appears in judicial opinions. One possibility is that judges are more likely to explicitly find that this factor disfavors a likelihood of confusion in cases in which they find an overall likelihood of confusion, evidence of the bi-directionality of coherence-based reasoning.<sup>108</sup> A second possibility is that judges are, in fact, biased in favor of finding no likelihood of confusion if there is no showing of actual confusion, despite their explicit recognition that a finding of actual confusion is both unnecessary and very difficult to prove.

*Presentation of Survey Evidence*

¶64 Our findings were consistent with Beebe’s observation that survey data is less frequently employed than one might expect given the conventional wisdom that survey evidence is routinely employed to prove a likelihood of confusion.<sup>109</sup> Of 139 preliminary injunction and bench trial opinions, only thirty opinions involved the presentation of survey evidence from either party (seven of these opinions involved presentation of survey evidence by both parties).

¶65 Also consistent with Beebe’s data was the relative infrequency with which courts credited survey evidence (only 32.4% of all survey evidence presented was credited). However, our data suggested a stronger relationship between credited survey evidence and the outcome of the likelihood of confusion test (91.7% of the opinions crediting the survey evidence also found in favor of the party presenting the survey), in contrast with Beebe’s results. It is unclear whether the credited survey evidence actually influenced the outcome of the analysis, or whether the outcome of the analysis made it more likely that the survey evidence would be credited. The following sections discuss these results in more detail.

---

<sup>105</sup> *The Sports Authority, Inc. v. Prime Hospitality Corp.*, 89 F.3d 955, 963 (2d Cir. 2006) (quoting *W.W.W. Pharm. Co., Inc. v. Gillette Co.*, 984 F.2d 567 (2d Cir. 1993)).

<sup>106</sup> *E.g., Maternally Yours, Inc. v. Your Maternity Shop, Inc.*, 234 F.2d 538, 542 (2d Cir. 1956).

<sup>107</sup> *See, e.g., Lois Sportswear, U.S.A., Inc. v. Levi Strauss & Co.*, 799 F.2d 867, 875 (2d Cir. 1986) (citing *W.E. Bassett Co. v. Revlon, Inc.*, 435 F.2d 656, 662 (2d Cir. 1970)) (“[A]ctual confusion is very difficult to prove and the [Lanham] Act requires only a likelihood of confusion as to source.”).

<sup>108</sup> See *supra* notes 29-31 and accompanying text.

<sup>109</sup> Beebe, *supra* note 1, at 1641.

*Plaintiff Presentation of Survey Evidence*

¶66 Plaintiffs presented survey evidence in twenty-four of 139 preliminary injunction and bench trial opinions (17.3%), but this evidence was favorably credited by the court in only one-quarter, or six of these opinions. When the survey was credited, however, the court always found that the actual confusion factor favored the plaintiff and always ultimately found a likelihood of confusion (a plaintiff win rate of 100%). In the eighteen opinions where the survey was not credited, none of the plaintiffs were able to win the actual confusion factor, and only three were nevertheless able to obtain a finding of a likelihood of confusion (a plaintiff win rate of 16.7%).

*Adverse Inference Against Plaintiff for Failure to Present Survey Evidence*

¶67 Despite the overall infrequency of opinions in which plaintiffs actually presented survey evidence, courts drew an adverse inference against the plaintiff for failure to submit survey evidence showing actual confusion in thirteen opinions of our set of 139 preliminary injunction and bench trial opinions. Courts that drew an adverse inference against the plaintiff for failure to present survey evidence were less likely both to find that this factor favored a likelihood of confusion and to find an overall likelihood of confusion. These data are particularly remarkable given that five of these thirteen opinions were preliminary injunction opinions, in which one would be less likely to expect the parties to be able to have had the time to prepare credible survey evidence.

¶68 Of the thirteen cases in which the adverse inference was drawn, only one resulted in a finding that this factor favored a likelihood of confusion (7.7%),<sup>110</sup> and only two opinions resulted in an overall finding of a likelihood of confusion (a plaintiff win rate of 15.4%). Of the 102 preliminary injunction and bench trial opinions where the lack of plaintiff survey evidence did not draw an explicit adverse inference from the judge, thirty-two resulted in a finding that this factor favored a likelihood of confusion (31.4%), and fifty-seven resulted in an overall finding of a likelihood of confusion (a plaintiff win rate of 55.9%).

*Defendant Presentation of Survey Evidence*

¶69 Defendants presented survey evidence in thirteen of 139 preliminary injunction and bench trial opinions (9.4%), and the court favorably credited this evidence in six of the thirteen opinions (46.2%). Defendants who presented credited survey evidence enjoyed a much higher likelihood of a finding of no likelihood of confusion, with only one of the six facing a finding of a likelihood of confusion (a defendant win rate of 83%). Of the seven opinions that did not credit the defendants' survey evidence, four found a likelihood of confusion (a defendant win rate of 57.1%).

*Factor 6: Intent*

¶70 The intent factor considers "whether the defendant adopted its mark with the intention of capitalizing on plaintiff's reputation and goodwill and any confusion between his and the senior user's product."<sup>111</sup> In our five-node classification tree, a finding that the intent factor favored confusion led to an overall finding of a likelihood of confusion in all of the eighty-six opinions in which the similarity factor favored a likelihood of confusion. However, a finding that this factor favored confusion was neither sufficient nor necessary to demonstrate an overall likelihood of confusion: 11.6% of opinions in which the court found bad faith did not result in a finding of a likelihood of confusion, while 20.8% of opinions found a likelihood of confusion despite a finding of no bad faith.

¶71 Of the preliminary injunction and bench trial opinions only, a finding that the intent factor favors a likelihood of confusion correlates with an ultimate finding of a likelihood of confusion with a value of 0.5282.

---

<sup>110</sup> Casa Editrice Bonechi, S.R.L. v. Irving Weisdorf & Co., 1995 U.S. Dist. LEXIS 12849 (S.D.N.Y. Sept. 6, 1995). The plaintiff seems to have overcome this adverse inference through the presentation of non-survey evidence such as witness testimony of retailer confusion. *Id.* at \*29-30.

<sup>111</sup> W.W.W. Pharm. Co., Inc. v. Gillette Co., 984 F.2d 567, 575 (2d Cir 1993) (quoting Lang v. Ret. Living Pub. Co., 949 F.2d 576, 583 (2d Cir. 1991)).

Conversely, a finding that this factor militates against a likelihood of confusion correlates with an ultimate finding of no likelihood of confusion with a value of 0.5825.

¶72 These findings are consistent with Beebe’s observations that the intent factor is highly relevant to the outcome of the multifactor test, although our data revealed a slightly looser correlation than Beebe’s. Beebe found that a finding of bad faith intent did not result in a finding of a likelihood of confusion in only two of the bench trial and preliminary injunction opinions in his sample of 192 such opinions, whereas we found five such instances in our sample of 139 preliminary injunction and bench trial opinions.<sup>112</sup> In these five opinions, a finding that the similarity factor did not favor a likelihood of confusion seems to have trumped the bad faith of the defendant. Our findings are also consistent with the Second Circuit holding that a finding of bad faith intent raises a “rebuttable legal presumption that the actor’s intent to confuse will be successful.”<sup>113</sup>

*Factor 7: The Quality of Defendant’s Mark*

*The Problem with Quality*

¶73 Beebe calls the quality factor “an embarrassment to the multifactor test.”<sup>114</sup> Empirically, quality played a very small role—perhaps a useless role—in the multifactor test. In 116 out of the total 206 cases, courts did not even make a quality determination. For the preliminary injunction/bench trial cases—*i.e.*, those cases in which the plaintiff either ultimately won or lost and there were not any outstanding factual issues, a quality determination was not made in seventy-six of the 139 cases. When courts made such a finding, it had little predictive value. A finding that quality favored a likelihood of confusion, or disfavored a likelihood of confusion, had the lowest correlation with win and loss rates of any of the factors. Further, when quality was found to favor a likelihood of confusion, those cases had much higher stampede scores, which might indicate that a finding favoring a likelihood of confusion determination was made not because of consideration of that factor, but because of stampeding.

¶74 Quality is inherently problematic in that it is quite subjective. Aside from that, however, the biggest problem lies in the courts’ use of two mutually contradictory theories of the relevant test for quality. Problems arise when different courts subscribe to different theories, as well as when individual courts subscribe to both theories—ignoring or seeming unaware of the conflict. When courts employ the most frequently used theory, tarnishment, the fact that the junior product is of inferior quality favors a likelihood of confusion.<sup>115</sup> When they apply the less popular theory, similarity, the fact that the junior product is of similar quality as the senior product favors a likelihood of confusion.<sup>116</sup> Some courts acknowledge that there are two theories and that they conflict, but nonetheless inexplicably determine that tarnishment should govern.<sup>117</sup> Still, other courts acknowledge there are two theories, but fail to address the fact that they conflict, leading to a “heads I win, tails you lose” situation:

---

<sup>112</sup> Beebe, *supra* note 1, at 1628-29.

<sup>113</sup> *Samara Bros. v. Wal-Mart Stores, Inc.*, 165 F.3d 120, 127 (2d Cir. 1998), *rev’d on other grounds*, *Wal-Mart Stores, Inc. v. Samara Bros., Inc.*, 529 U.S. 205 (2000).

<sup>114</sup> Beebe, *supra* note 1, at 1644-45.

<sup>115</sup> *See Conopco, Inc. v. Cosmair, Inc.*, 49 F. Supp. 2d 242, 255 (S.D.N.Y. 1999) (“This factor is primarily concerned with whether the senior user’s reputation could be jeopardized by virtue of the fact that the junior user’s product is of inferior quality.”).

<sup>116</sup> *Landscape Forms, Inc. v. Columbia Cascade Co.*, 117 F. Supp. 2d 360, 367 (S.D.N.Y. 2000) (“There is also little dispute that the two product lines are of similar quality—a factor also weighing in plaintiff’s favor.”).

<sup>117</sup> *E.g., Clinique Labs., Inc. v. Dep Corp.*, 945 F. Supp. 547, 555-56 (S.D.N.Y. 1996) (citations omitted) (“Although Clinique has not presented evidence of the Basique products’ inferior quality, courts in this Circuit have held that a ‘lack of marked difference in quality between goods supports the inference that they emanate from the same source.’ At the same time, however, the quality factor protects the senior user’s ‘good reputation associated with [its] mark from the possibility of being tarnished by inferior merchandise.’ These two doctrines seem to conflict: goods of either equivalent or inferior quality can lead to a likelihood of confusion with the senior user’s product. While the former interpretation more directly addresses the likelihood of consumer confusion, the latter emphasizes precisely why a likelihood of confusion is damaging to the senior user. Considering that Dep currently represents that Basique is as good as Clinique, and mindful of the fact that Clinique cannot control the quality of Basique so that its reputation could be damaged if Basique products turn out to be inferior, I conclude that this factor weighs in favor of a likelihood of confusion.”); *Revlon Consumer Prods. Corp. v. Jennifer Leather Broadway, Inc.*, 858 F. Supp. 1268, 1275 (S.D.N.Y. 1994) (citations omitted) (“Plaintiff does not contend that defendants’ services are not of a high quality; plaintiff merely contends that it has no ability to control the quality of defendant’s services. To this issue, the Second Circuit has noted the potential irony in applying the ‘quality’ factor within a likelihood of confusion analysis; to wit, if the product or service is of an equally high

The analysis of the quality of the defendants' product "is primarily concerned with whether the senior user's reputation could be jeopardized by virtue of the fact that the junior user's product is of inferior quality." In addition, if the plaintiff's goods and the defendants' goods are of comparable quality, that fact may increase the likelihood of confusion. The magazines at issue here are of comparable quality in terms of their paper, printing, and binding. Therefore, the plaintiff's reputation is not harmed by the existence of the defendants' use of an "O" on their publication. However, the similarity in quality between the publications may somewhat increase the likelihood of confusion. Thus, this factor favors the plaintiff only slightly, if at all.<sup>118</sup>

¶75 Judge Friendly gave little guidance on how the quality factor was to be tested, as he simply noted that "the quality of defendant's product" was a factor.<sup>119</sup> In early Southern District cases adopting the *Polaroid* test, courts kept their analysis of the quality factor vague, though they indicated a preference for the tarnishment theory.<sup>120</sup> This preference was demonstrated in our data set as well.<sup>121</sup> In our all-cases dataset, courts favor the tarnishment theory alone over a combination of both theories almost three to one. Courts subscribe only to the tarnishment theory in eighty-seven of 206 cases (sixty of 139 in the PI/BT dataset). Courts adhere only to the similarity theory in twelve of 206 cases (or nine of 129 in the PI/BT). In thirty-three of 206 cases (and twenty-two of 139 in the PI/BT), courts mention both theories.

#### *Should Quality Remain a Factor?*

¶76 Beebe argues that the quality test is an embarrassment on a number of grounds—first, tarnishment should not be relevant to a finding of a likelihood of confusion, second, because similarity is already "addressed under the proximity factor," and third, most strongly, "because the factor is so utterly pliable."<sup>122</sup> It seems then that even if courts consistently subscribed to only one theory, Beebe would criticize the courts' approach. But the use of both theories, applied independently and even simultaneously, exacerbates the issue. Because there are two theories applicable to the quality factor in the Second Circuit, and because in some senses these two theories are opposite questions, a judge could pick whichever theory he or she preferred and find a likelihood of confusion. This makes the factor even more pliable than it naturally is, given the inherently subjective nature of a quality determination.

¶77 However, there is perhaps some hope for the quality factor, if it remains (though ideally we agree with Beebe's suggestion that it be removed). Over time there seems to be a slight tendency to limit explicit use of the tarnishment test.<sup>123</sup> Courts have become less likely in recent years to expressly adopt any theory when analyzing the quality of the products factor. While this alone is not a cause for cheer, in particular, courts appear to be less explicit in endorsing the generally dominant tarnishment theory. In the years 2006 to 2008 in our all cases dataset, for example, only twelve of thirty three cases were found to mention the tarnishment theory, while the corresponding figure in 1993 to 1996 was twenty one of forty five. Our PI/BT dataset revealed a similar trend, and the corresponding numbers for 2006 to 2008, and 1993 to 1996, were seven of

---

standard of quality this might lead to increased confusion as to its source. Plaintiff reads too much into this language. Read literally it would create a per se rule that whenever a defendant in a Lanham Act case produces a product of equal quality to the plaintiff's product, there is a greater likelihood of customer confusion. That cannot be what the Second Circuit intended . . .").

<sup>118</sup> *Brockmeyer v. The Hearst Corp.*, 2002 U.S. Dist. LEXIS 11725 (S.D.N.Y. June 22, 2002).

<sup>119</sup> *Polaroid Corp. v. Polaroid Elecs., Corp.*, 287 F.2d 492, 495 (2d Cir. 1961).

<sup>120</sup> *See, e.g., Field Enterprises Ed. Corp. v. Grosset & Dunlap, Inc.*, 256 F. Supp 382, 390 (S.D.N.Y. 1966) (indicating that any likelihood of confusion was de minimus and stating "[f]inally, although defendant's series is moderately priced, the content of their books is of good quality . . ."); *Blue Bell, Inc. v. Jaymar-Ruby, Inc.*, 179 U.S.P.Q. 665, 666 (S.D.N.Y. 1973) (finding no infringement and noting regarding quality that there was "no evidence as to any lack in the quality of defendant's product").

Even an early case, however, while subscribing to the tarnishment theory, indicates its ridiculousness by noting that the lack of tarnishment could still lead to confusion:

There is no challenge to the quality of defendant's product. Defendant is an established manufacturer of pharmaceuticals in a closely regulated industry, and plaintiff does not contend that use of the Vagestrol mark on defendant's suppositories has in any way harmed, or will harm, its own or its product's reputation, except to the extent that fear of confusion of names on the part of a pharmacist might deter a physician from prescribing Vagitrol.

*Syntex Labs., Inc. v. Norwich Pharmacal Co.*, 315 F. Supp. 45, 52 (S.D.N.Y. 1970).

<sup>121</sup> Note that in both datasets, some courts do not explain which theory they are relying on when analyzing this factor, accounting for the fact that the numbers for each theory do not add up to the overall number of cases in the dataset.

<sup>122</sup> Beebe, *supra* note 1, at 1645.

<sup>123</sup> *See supra* p. 24.

twenty-one, and twenty of thirty-nine, respectively. While the similarity theory of the quality factor may be redundant in a test that already employs a proximity factor, it is at least true that if the quality of products is similar, this will probably result in an increased likelihood of confusion. Moreover, the use of this version of the factor, so long as courts are aware of its double-counting effect, will at least not distort the result of the *Polaroid* analysis.

*Factor 8: The Sophistication of Consumers*

¶78 Sophistication is not a particularly useful factor in terms of predicting outcomes. Of the PI/BT cases only, a finding that the consumer sophistication factor favors the likelihood of confusion correlates with an ultimate finding of a likelihood of confusion with a value of 0.4085. Conversely, a finding that this factor militates against a likelihood of confusion correlates with an ultimate finding of no likelihood of confusion with a value of 0.4461. Of all the cases, a finding that the consumer sophistication factor favors the likelihood of confusion correlates with an ultimate finding of a likelihood of confusion with a value of 0.3633. Conversely, a finding that this factor militates against a likelihood of confusion correlates with an ultimate finding of no likelihood of confusion with a value of 0.4998. Given that it does not seem to be used to great effect in the test, the sophistication factor could be eliminated from the analysis and there would likely be no change in the results. This seems an odd result, because, as Beebe argues,<sup>124</sup> the sophistication of consumers appears highly relevant to determining whether there is a likelihood of confusion. Perhaps courts' heavier reliance on the sophistication factor in the future would lead to better applications of the *Polaroid* test.

PART III: OTHER CONSIDERATIONS

*Competition*

¶79 One new variable we coded for that Beebe did not was whether the products were in competition. We were curious to see if it would affect and/or predict win rates, and how it would relate to the proximity factor.<sup>125</sup> Tables D and E set out plaintiff win rates by whether products were in competition and by factor.

¶80 We found that if products were non-competing, there was, as expected, a much lower win rate than the average win rate. However, knowing if products were competing or not competing was generally less useful at predicting outcome than knowing any other factor except quality of the products. Of the preliminary injunction/bench trial cases, a finding that the products are competing correlates with an ultimate finding of a likelihood of confusion with a value of 0.2677. Conversely, a finding that the products are not competing correlates with an ultimate finding of no likelihood of confusion with a value of 0.2907. For all the cases, a finding that the products are competing correlates with an ultimate finding of a likelihood of confusion with a value of 0.3128. Conversely, a finding that the products are not competing correlates with an ultimate finding of no likelihood of confusion with a value of 0.4129.

¶81 However, if products were competing, knowing the outcome of the “Actual Confusion” test was quite predictive. If products were competing, evidence of actual confusion became extremely important—the most useful factor for a plaintiff to win. When evidence of actual confusion favored a likelihood of confusion and products were competing, only one case came out against a likelihood of confusion.

¶82 Competition was, as anticipated, correlated with proximity, although it was not one-to-one. A fuller discussion of the relationship between competition and proximity may be found in the section on the proximity factor above.

---

<sup>124</sup> “Probably the only non-core factor that deserves to be in the multifactor test is the consumer sophistication factor. It makes sense, and has been confirmed empirically, that the more sophisticated the consumers, the more care with which they will treat their search and purchasing decisions.” Beebe, *supra* note 1, at 1642.

<sup>125</sup> See Appendix, Table 10.

*“Any One Factor May Be Dispositive”*

¶83 Although Second Circuit precedent cautions that no single *Polaroid* factor is dispositive,<sup>126</sup> eight of the 206 opinions in our sample stated that any one factor may be dispositive.<sup>127</sup> Interestingly, these eight opinions did not yield a higher plaintiff win rate, with only two resulting in a finding of a likelihood of confusion. Only two of the opinions elaborated on this point, and both clarified that it was the similarity factor that could be dispositive if the marks were sufficiently dissimilar, consistent with our finding that winning the similarity factor was necessary to win the likelihood of confusion analysis. In both opinions, the court found that the similarity factor did not favor a likelihood of confusion and ultimately concluded that there was no likelihood of confusion.<sup>128</sup>

*Party Names*

¶84 Another inquiry we made that Beebe did not was to what extent a finding that a mark contained the proper name of its owner (or a member of its owner’s family) affected the ultimate finding of a likelihood of confusion. Such a finding has no effect on the ultimate likelihood of confusion. There were only nineteen cases in which the court mentioned a mark’s origins in the owner’s name. No trends were discernable in those cases, as the court found no ultimate likelihood of confusion in a majority of the six cases where only the plaintiff’s mark came from the owner’s name, nor in a majority of the five cases where only the defendant’s mark came from the owner’s name.<sup>129</sup> There were no statistically significant correlations between judicial notice that a mark was named for its owner and an ultimate finding in favor of that owner. The only discernable policy conclusion one can draw from this inquiry is support for the proposition that a trademark owner should choose a fanciful or arbitrary name for his mark.<sup>130</sup> A plea that one’s mark was named for oneself or one’s family will be ineffectual in front of a judge.

*Additional Factors*

¶85 In addition to the *Polaroid* factors, a few opinions included consideration of other additional factors in the likelihood of confusion analysis. Some examples of additional factors included: presence of initial interest confusion (potential consumers are attracted to the junior mark because of its similarity to the senior mark, even though there is no confusion at the time of purchase)<sup>131</sup> and the use of comparative advertising and disclaimers.<sup>132</sup>

¶86 These factors, however, were almost always fact-specific, and it is difficult to generalize about the impact of any one individual additional factor. The following conclusion, therefore, is limited to the impact of these additional factors as a whole. From our PI/BT dataset of 139 cases, there were seventeen opinions (12.2%) that considered some additional factor. In these opinions, the plaintiff win rate jumps to about 70%, whereas those opinions that do not consider at least one additional factor yield a plaintiff win rate of 46%. Granted, the sample size is extremely limited, but if these results are reflective of actual trends, it appears that considering additional factors may have a significant impact on the courts’ analysis.

¶87 One possible explanation for this effect is that those cases that lead courts to consider and analyze some salient additional factor are simply those cases where the court feels that some additional factor is dispositive

---

<sup>126</sup> *E.g.*, *Louis Vuitton Malletier v. Dooney & Bourke, Inc.*, 454 F.3d 108, 118 (2d Cir. 2006) (citing *Brennan’s, Inc. v. Brennan’s Rest., L.L.C.*, 360 F.3d 125, 130 (2d Cir. 2004)).

<sup>127</sup> We cannot compare our results to Professor Beebe’s as we do not believe he collected data on this variable.

<sup>128</sup> *See Playtex Prods. Inc. v. Georgia-Pacific Inc.*, 2003 U.S. Dist. LEXIS 13981 (S.D.N.Y. Aug. 12, 2003); *Kaufman & Fisher Wish Co. v. FAO Schwartz*, 184 F. Supp. 2d 311 (S.D.N.Y. 2001).

<sup>129</sup> In the other eight cases, the court noticed that both the plaintiff’s mark and the defendant’s mark were based on their respective proper names. These give little insight into whether basing a mark on your name helps win your case, as both parties would seem equally entitled to any extra benefits.

<sup>130</sup> *See supra* pp. 25-26.

<sup>131</sup> *Katz v. Modiri*, 283 F. Supp. 2d 883, 899 (S.D.N.Y. 2003); *New York State Soc. of Certified Public Accountants v. Eric Louis Assocs., Inc.*, 79 F. Supp. 2d 331, 342 (S.D.N.Y. 1999).

<sup>132</sup> *Clinique Lab., Inc. v. Dep Corp.*, 945 F. Supp. 547, 556 (S.D.N.Y. 1996).

in finding a likelihood of confusion. That is, when deciding the question of the likelihood of confusion, if a court does not feel that the additional factor is important or dispositive, it is intuitively probable that the court would not mention the additional factor at all. Conversely, if a court feels strongly enough about an additional factor so that it analyzes it in the opinion, it is conceivable that the court will rely heavily on that factor when deciding the ultimate issue of a likelihood of confusion. Another possibility, in accordance with the coherence-based reasoning model, is that courts distinguish and discuss the additional factor in order to justify a finding of a likelihood of confusion.<sup>133</sup>

## CONCLUSION

¶88 Our results of studying the fifteen-year set of cases in the District Court for the Southern District of New York were, for the most part, consistent with Beebe's national study of 2001-2004. There seems to only be a small number of key factors, with similarity being paramount among them. We noted less stampeding than Beebe observed, although he too found less stampeding in the district courts of the Second Circuit. We did observe what may be a moderate decline in the rate of finding a likelihood of confusion as well as a potential decline in the use of the tarnishment theory alone. We also observed an apparent decrease in the importance of inherent strength and its connection to commercial strength. Of the new variables we observed, the one of most interest may be the consideration by the court of additional factors, as those seemed to affect decision-making the most. We would suggest a reformed test along the lines of the one Beebe recommends—focusing on similarity, proximity as perceived by the relevant consumers, evidence of actual confusion, marketplace strength, and the infringer's intent.<sup>134</sup>

¶89 Our recommendations for further study include going back further in time regarding the quality factor and/or seeing how this factor works in other circuits, determining if more factors were carefully considered when the test was first introduced (considering the historical development of rational core attributes heuristics), and looking at what additional factors are considered in a larger data set to see if they have the same effects. These would even better aid us to prescribe a new trademark test, as the current one, while not broken, could certainly use some help.

---

<sup>133</sup> See *supra* text accompanying note 33.

<sup>134</sup> Beebe, *supra* note 1, at 1646-47.

APPENDIX

**Table 1:** *Coding Rules and Explanations of Variables*

- (1) The opinion should be excluded if the case falls into any of the following categories:
  - Counterfeiting opinions
  - Opinions that involve a breach of a franchising, licensing, or distribution agreement
  - Opinions on motion to dismiss
  - Opinions where the non-moving party failed to appear
  - Opinions where the outcome of the multifactor test is reversed on appeal

If you find such an opinion, contact your coding partner, who should confirm independently. Retain and code reverse confusion opinions and fact patterns in which  $\Delta$  repackaged  $\pi$ 's goods.
- (2) If your opinion includes more than one *Polaroid* analysis (e.g., multiple marks), contact your coding partner to coordinate separate entries for each analysis.
- (3) **Explicit statements from the judge always trump for the purposes of coding, including preemption of the below rules** (e.g., if the judge says that marks are slightly similar, but then notes that this factor weighs against a finding of confusion, code it as found not to favor confusion).
- (4) Without explicit statements to the contrary from the judge, when a judge states that a factor “slightly” or “moderately” favors a finding of confusion, code as found to favor confusion.
- (5) A judge’s explicit statements about the valence of factors in the “summing up” section of his or her *Polaroid* analysis trumps any previous statements he or she has made.
- (6) *Abercrombie* determinations should be coded regardless of where they are located in the opinion.
- (7) If a judge makes an *Abercrombie* determination that the mark is fanciful, arbitrary, or suggestive, code “1” for “f1inhstr” (mark is inherently strong); if he or she determines it is descriptive or generic, code “1” for “f1inhweak” (mark is inherently weak).
- (8) Surveys:
  - (a) if the judge states that a survey had some problems but does not explicitly invalidate the survey, code “1” in the “other” category of the party that offered the survey;
  - (b) if the judge credits a survey but states that it weighs in favor of the opposing party, code “1” in the “other” credit category of the party that offered the survey;
  - (c) if a party offers multiple surveys but not all are credited, still code “1” in the credited column.
- (9) If the judge states in the bridging the gap factor that there is “no gap to bridge” (or equivalent), code it as “other” in the absence of explicit guidance from the court.
- (10) If the judge notes that actual confusion can be dispositive by itself, in addition to coding “1” for “actualdisp,” code “1” for “any1disp” (any one factor can be dispositive).

Table 1

Year	Year	Year opinion was filed	YYYY
<b>Posture</b>	Pi	Preliminary injunction	0/1
	Bt	Bench trial	0/1
	Sjp	Summary judgment motion by plaintiff, none by defendant	0/1
	Sjd	Summary judgment motion by defendant, none by plaintiff	0/1
	Sjc	Cross summary judgment motion	0/1
	Tro	TRO, but not PI	0/1
	Mjrr	Mag. Judge Report and Recommendation	0/1
	Mtd	Motion to dismiss	0/1
<b>General descriptors</b>	Tm	Case involved trademark infringement	0/1
	Td	Case involved trade dress infringement	0/1
<b>Dispositive Factors</b>	any1disp	Court explicitly states that any one <i>Polaroid</i> factor can be dispositive	0/1
	actualdisp	Court explicitly states that actual confusion can be dispositive	0/1
<b>Factor 1: The Strength of the Plaintiff's Mark</b>	f1conf	Found to favor confusion	0/1
	f1noconf	Found not to favor confusion	0/1
	f1other	Found to be neutral/unclear/other	0/1
	Abfanonly	Mark held to be fanciful and nothing else	0/1
	Abaronly	Mark held to be arbitrary and nothing else	0/1
	Absugonly	Mark held to be suggestive and nothing else	0/1
	Abdesonly	Mark held to be descriptive and nothing else	
	Abgenonly	Mark held to be generic and nothing else	0/1
	Abother	<i>Abercrombie</i> class unclear/other	0/1
	f1inhstr	Mark is inherently strong	0/1
	f1inhweak	Mark is inherently weak	0/1
	f1inhunclear	Inherent strength unclear	0/1
	f1commstr	Mark is commercially strong (strong secondary/acquired meaning)	0/1
f1commweak	Mark is commercially weak (weak showing of secondary/acquired meaning)	0/1	
f1communclear	Commercial strength unclear	0/1	
<b>Factor 2: The Similarity of Plaintiff's and Defendant's Marks</b>	f2conf	Found to favor confusion	0/1
	f2noconf	Found not to favor confusion	0/1
	f2other	Found to be neutral/unclear/other	0/1
<b>Factor 3: The Proximity of Plaintiff's and Defendant's Products</b>	f3conf	Found to favor confusion	0/1
	f3noconf	Found not to favor confusion	0/1
	f3other	Found to be neutral/unclear/other	0/1
<b>Factor 4: Evidence of</b>	f4conf	Found to favor confusion	0/1

<b>Bridging Gap</b>	f4noconf	Found to not favor confusion	0/1
	f4other	Found to be neutral/unclear/other	0/1
	f4bridge	Found will bridge, to favor confusion	0/1
	f4nogap	Found no gap to bridge	0/1
<b>Factor 5: Evidence of Actual Confusion</b>	f5conf	Found to favor confusion	0/1
	f5noconf	Found not to favor confusion	0/1
	f5other	Found to be neutral/unclear/other	0/1
	f5advinf	Court makes adverse inference for plaintiff's failure to present survey evidence	0/1
	f5psurv	Plaintiff presented survey evidence	0/1
	f5psurvc	Plaintiff's survey evidence was credited	0/1
	f5psurvcother	Mixed/unclear credit given to $\pi$ 's survey (or favored $\Delta$ )	0/1
	f5dsurv	Defendant presented survey evidence	0/1
f5dsurvc	Defendant's survey evidence was credited	0/1	
f5psurvcother	Mixed/unclear credit given to $\Delta$ 's survey (or favored $\pi$ )	0/1	
<b>Factor 6: Defendant's Intent</b>			
	f6conf	Found to favor confusion	0/1
	f6noconf	Found not to favor confusion	0/1
	f6other	Found to be neutral/unclear/other	0/1
<b>Factor 7: The Quality of Defendant's Goods</b>			
	f7conf	Found to favor confusion	0/1
	f7noconf	Found not to favor confusion	0/1
	f7other	Found to be neutral/unclear/other	0/1
	f7tarn	Theory is that defendant's low quality harms plaintiff and supports finding of a likelihood of confusion	0/1
	f7sim	Theory is that similar quality of goods supports finding of a likelihood of confusion	0/1
<b>Factor 8: The Sophistication of the Consumers</b>			
	f8conf	Found to favor confusion	0/1
	f8noconf	Found not to favor confusion	0/1
	f8other	Found to be neutral/unclear/other	0/1
<b>Additional Factors</b>	Addfact	Court considered factor(s) in addition to <i>Polaroid</i> factors	0/1
<b>Outcome</b>			
	Conf	A likelihood of confusion is found	0/1
	Noconf	A likelihood of confusion is not found	0/1
	Fi	Found to be outstanding issues of fact	0/1
<b>Factor 9: Competition</b>			
	f9comp	Found to be competing (presumption if found no gap in F5)	0/1
	f8nocomp	Found not to be not competing (presumption if found the existence of a gap in F5)	0/1
	f8other	Found to be neutral/unclear/other	0/1
<b>Factor 10: Party Name</b>			
	f8pname	Found plaintiff's name in mark	0/1
	f8dname	Found defendant's name in mark	0/1

Figure I: Next Best One-Node Classification Trees

SECOND AND THIRD BEST ONE-NODE CLASSIFICATION TREES

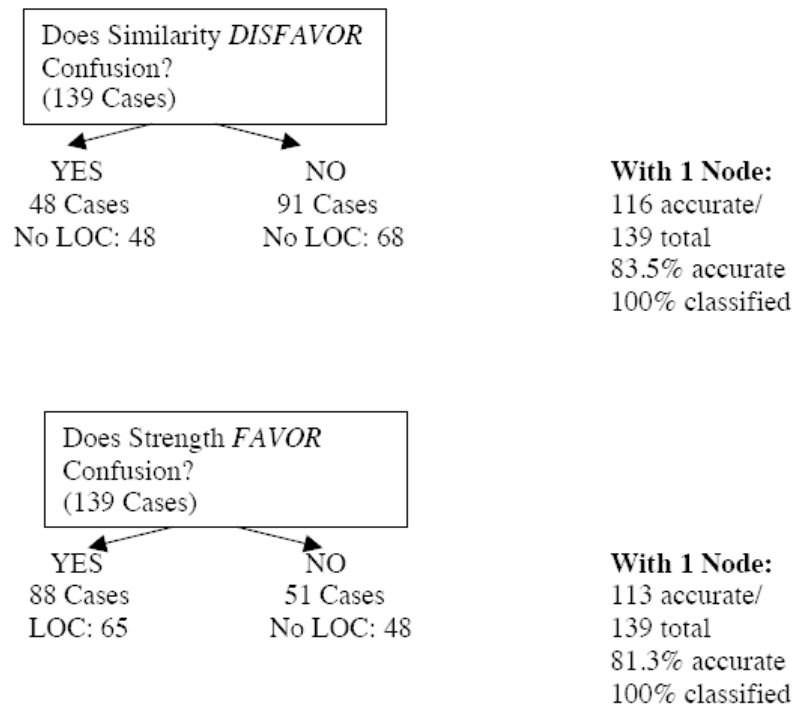


Figure II: Next Best Two-Node Classification Trees

SECOND AND THIRD BEST TWO-NODE CLASSIFICATION TREES

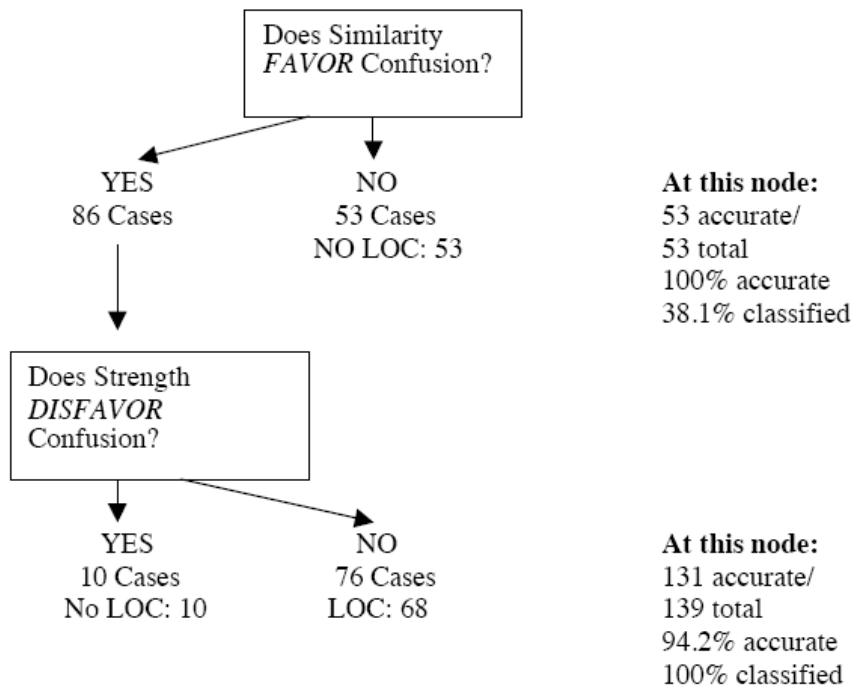


Figure III: Next Best Five-Node Classification Tree – Similarity

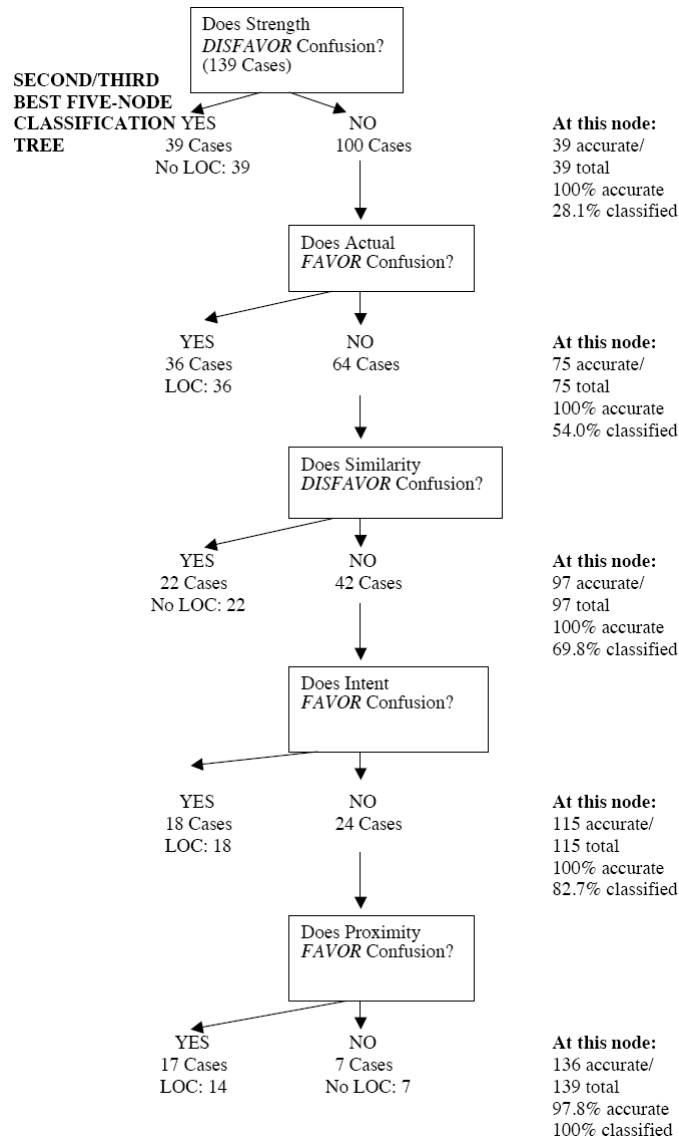
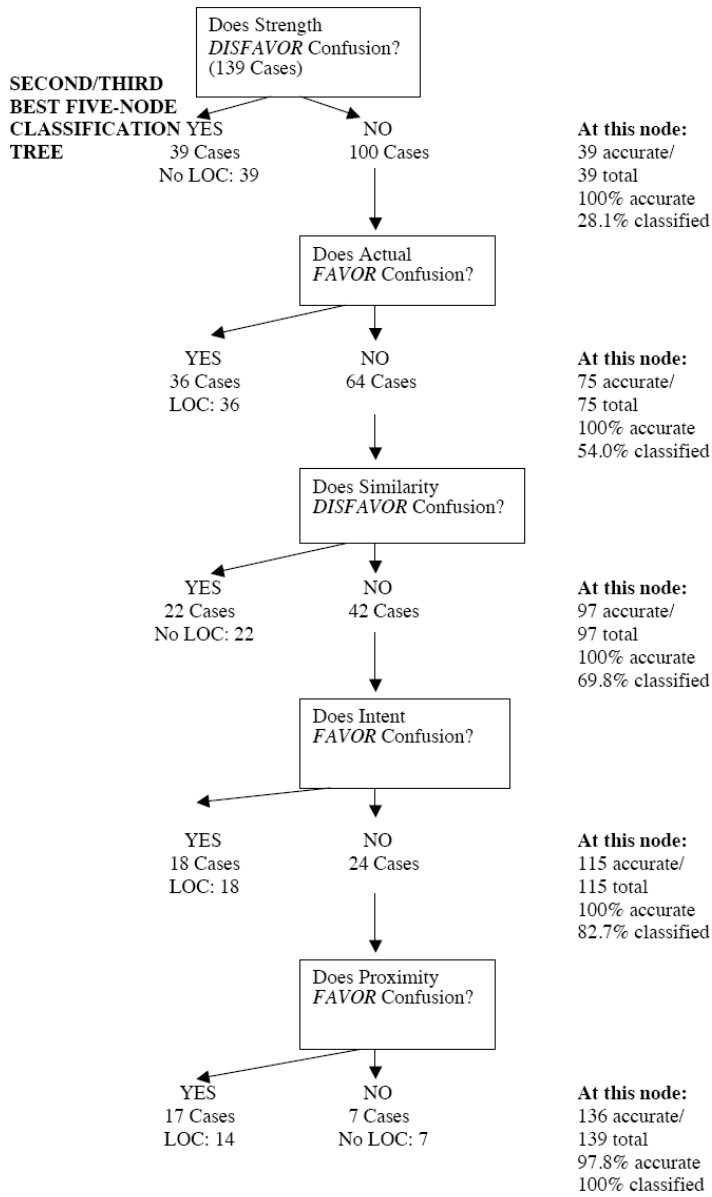


Figure IV: Next Best Five-Node Classification Tree – Strength



**Table 2:** *Additional Historical Trends Data*

Historical trends, all cases dataset	1993-1996	1997-1999	2000-2002	2003-2005	2006-2008	Total
1) # of cases	45	46	39	43	33	206
2) Ratio of a likelihood of confusion found	21:21:3	21:21:4	17:18:4	14:25:4	12:18:3	85:103:18
3) # of “any one factor may be dispositive” cases	0	1	3	2	2	8
4) # of similarity theory only for quality	1	3	1	2	5	12
5) # of tarnishment theory only for quality	21	25	16	13	12	87
6) # of both similarity and tarnishment theories	9	5	6	9	4	33
7) Fraction where no bridge favors LoC	5/16	13/28	6/12	5/18	0/7	29/81
8) Fraction where no bridge = neutral/irrelevant	11/16	15/28	6/12	13/18	7/7	52/81
9) Ratio inherently strong where F1 favors LoC	24:3:2	18:7:3	17:3:2	16:5:1	14:6:1	89:24:9
10) Ratio comm. strong where F1 favors LoC	18:0:2	17:0:0	15:1:0	22:0:0	15:1:0	87:2:0
11) Ratio inherently strong where LoC is found	21:7:1	15:11:2	13:8:1	10:10:2	10:10:1	69:46:7
12) Ratio commercially strong where LoC is found	13:5:2	16:1:0	9:5:2	13:7:2	9:7:0	60:23:6

Historical trends, BT/PI dataset	1994-1996	1997-1999	2000-2002	2003-2005	2006-2008	Total
1) # of cases	39	34	20	25	21	139
2) Ratio of a likelihood of confusion found	21:18:0	18:16:0	11:10:0	9:16:0	9:11:0	68:71:0
3) # of “any one factor may be dispositive” cases	0	1	2	0	1	4
4) # of similarity theory only for quality	1	2	1	1	4	9
5) # of tarnishment theory only for quality	20	17	9	7	7	60
6) # of both similarity and tarnishment theories	7	3	3	5	4	22
7) Fraction where no bridge favors LoC	5/14	10/22	5/8	2/10	0/5	22/59
8) Fraction where no bridge = neutral/irrelevant	9/14	12/22	3/8	8/10	5/5	37/59
9) Ratio inherently strong where F1 favors LoC	23:3:1	15:4:1	9:2:1	11:3:0	8:3:1	66:15:4
10) Ratio comm. strong where F1 favors LoC	16:0:1	16:0:0	7:0:0	12:0:0	10:1:0	61:1:1
11) Ratio inherently strong where LoC is found	20:7:0	12:8:0	8:5:0	8:6:0	7:5:0	55:31:0
12) Ratio commercially strong where LoC is found	13:4:0	15:1:0	6:1:0	8:4:0	6:5:0	48:15:0

When interpreting the tables, the following apply:

- Row 1 shows the number of cases in each 3-year basket.
- Row 2 shows the ratio of cases as to the ultimate finding of a likelihood of confusion where a likelihood of confusion is found, no likelihood of confusion is found, and outstanding issues of fact remain, respectively.
- Row 3 shows the number of cases in which the court included language that any one of the *Polaroid* factors may be dispositive in the likelihood of confusion analysis.
- Rows 4-6 show the number of cases where, in analyzing the quality factor, the court subscribes to a similarity theory only, a tarnishment theory only, or both theories,

respectively. The total of all three do not add up to the number of cases for that time period because some courts declined to subscribe to a particular theory.

- Rows 7 and 8 show the fraction of cases in which the court, analyzing the bridging the gap factor, treated the situation where there is no gap to bridge as favoring a likelihood of confusion on that factor, or is neutral or irrelevant, respectively. The denominator in each of those values reveal the total number of cases in a time period where the no gap to bridge situation arose, and the numerator reveals the court's treatment as to the absence of a gap to bridge.
- Rows 9 and 11 relate to the court's analysis of the inherent strength of the plaintiff's mark, in the strength of the mark factor. Row 9 reveals the ratio of cases where, given an inherently strong mark, the strength of the mark factor was found to favor confusion, militate against confusion, or is neutral or irrelevant, respectively. Row 11 reveals the ratio of cases where, given an inherently strong mark, an ultimate conclusion of a likelihood of confusion is found, a likelihood of confusion is not found, or outstanding issues of fact remain, respectively.
- Rows 10 and 12 relate to the court's analysis of the commercial strength of the plaintiff's mark in the strength of the mark factor. Row 10 reveals the ratio of cases where, given a commercially strong mark, the strength of the mark factor was found to favor confusion, militate against confusion, or is neutral or irrelevant, respectively. Row 12 reveals the ratio of cases where, given a commercially strong mark, an ultimate conclusion of a likelihood of confusion is found, a likelihood of confusion is not found, or outstanding issues of fact remain, respectively.

**Table 3:** *Proportion of Factor Outcomes Favoring a Likelihood of Confusion*

General Distinctiveness Classification	N	PW	Similarity of Marks	Proximity of Goods	Bridge Gap	Actual Confusion	Intent	Quality	Sophistication
Inherently Distinctive	1 2 2	.566	.623	.658	.312	.336	.336	.262	.328
Inherently Weak	4 2	.143	.309	.571	.262	.119	.119	.119	.214
Unclear	4 2	.262	.571	.476	.238	.191	.214	.167	.191
Total	2 0 6	.413	.549	.602	.286	.262	.267	.214	.277

**Table 4:** *Proportion of Factor Outcomes Favoring a Likelihood of Confusion*

Commercial Strength Classification	N	PW	Similarity of Marks	Proximity of Goods	Bridge Gap	Actual Confusion	Intent	Quality	Sophistication
Commercially Strong	91	.659	.747	.758	.385	.385	.385	.340	.374
Commercially Weak	53	.038	.245	.302	.151	.076	.076	.076	.113
Unclear	62	.371	.516	.629	.258	.242	.258	.145	.274
Total	20 6	.413	.549	.602	.286	.262	.267	.214	.277

**Table 5:** *Proportion of Factor Outcomes Favoring a Likelihood of Confusion*

Specific Abercrombie Classification	N	PW	Similarity of Marks	Proximity of Goods	Bridge Gap	Actual Confusion	Intent	Quality	Sophistication
Fanciful	3	.667	.667	.667	.333	.333	1.000	.667	.667
Arbitrary	23	.696	.783	.652	.435	.478	.261	.261	.304
Suggestive	48	.396	.458	.625	.313	.271	.229	.125	.229
Descriptive	27	.222	.482	.519	.296	.148	.185	.148	.222
Generic	2	.000	.000	1.000	.500	.000	.000	.000	.500
Total	103	.417	.534	.612	.334	.282	.243	.175	.262

**Table 6**

**All Cases Competition & Proximity Correlation:**

Competition?	N	%	Proximity?	N	Correlation
Competing	107	0.52	Prox LOC	95	0.888
			Prox No LOC	5	0.047
			Prox OTH	7	0.065
Non-Competing	68	0.33	Prox LOC	17	0.250
			Prox No LOC	44	0.647
			Prox OTH	7	0.103
Other Competing	31	0.15	Prox LOC	12	0.387
			Prox No LOC	5	0.161
			Prox OTH	14	0.452

Proximity?	N	%	Competition?	N	Correlation
Prox LOC	124	0.6	Competing	95	0.766
			Non-Competing	17	0.137
			Other Competing	12	0.097
Prox No LOC	54	0.26	Competing	5	0.093
			Non-Competing	44	0.815
			Other Competing	5	0.093
Prox OTH	28	0.14	Competing	7	0.250
			Non-Competing	7	0.250
			Other Competing	14	0.500

**Competition & Win Rate--PI/BT Cases**

	N	LOC	Win Rate
Overall	139	68	48.9%
Competing	74	45	60.8%
Non-Competing	46	13	28.3%