Stylistic Difference in Biblical Hebrew

E Bruce Brooks
University of Massachusetts at Amherst
(23 April 2018)

This article will introduce the BIRD test[1] as implemented for Biblical Hebrew. BIRD differs from other tests in that it is based not on content words, which *convey* message, but rather on high-frequency connectives, which *articulate* message; and that it measures, not single texts, but the *difference between two* texts.[2]

The 14 test words used for Biblical Hebrew include both independent and enclitic versions of concepts such as "and." In frequency order, the test words are:

ו	w'	"and"	0·1214	מן	min	"from"	0·0183	כי	ki	"for"	0·0108
ה	h'	"the"	0·0583	על	'al	"on"	0·0139	הב	b'ha-	"on the"	0·0106[3]
ל	l'	"to"	0·0410	אל	el	"to"	0·0133	הל	l'ha-	"to the"	0·0080[2]
ב	b'	"in"	0·0269	אשר	asher	"that"	0·0133	כ	k'	"as"	0·0057
את	et	[obj]	0·0265	לא	lo	"not"	0·0125				

The least common (כ) has the general frequency 0.0057. The smallest text for which כ would have E = 0·50 is 88 words. This is the recommended minimum sample size.

The Hebrew text here used is Leningradensis in the Michigan-Claremont lemmatized version. Some adjustments have been thought desirable:

• **Ketib/Qere** notations in the Masoretic text identify received and corrected readings; only the readings preferred by the Masoretic scholars have been retained.

• **Aramaic** text (as in Ezra and Daniel) has been excluded.

• **Headnotes** and other performance specifications in the Psalms (such as Selah) are not properly part of the text, and are bracketed off for purposes of stylistic analysis.

These changes reduce the working size of the corpus from 413,969 to 413,101 words. The frequencies given above are based on this adjusted text.

The 03c Septuagint translation sometimes reflects a Hebrew text superior to the Masoretic (6c). Those readings are not here followed.

---

[1]BIRD (**B**rooks **I**ndex of **R**hetorical **D**ifference) was devised by Bruce and Taeko Brooks for English and classical Chinese, and was implemented for Biblical Hebrew by Keith L Yoder, who has provided the computations here relied on. For the formulas, see Brooks **Chinese**.

[2]As in tennis, where player A can beat B, and B can beat C, but C can also beat A, it does not follow that if text A agrees with B, and B in turn with C, that A and C will closely agree.

[3]These test words combine a preposition with a following article.

Alpha v3 (2022)

Difference (D) is expressed by a number. The **Normal** range, from **0·51** to **0·75**, is neither very like nor very unlike. Such numbers occur between chapters of a work, or passages in the same area of discourse. We moderns can simulate a Biblical tone by borrowing King James' English, and the Biblical writers themselves could easily write in a "Biblical" style; our **Normal**. That relationship is thus not very informative.

Low readings, from **0·50** down, indicate significant closeness of style, as between parts of a continuous narrative. They also occur when a writer has another text in mind, as when the end of a sermon harks back to its beginning; a "lookback." An interpolator may successfully imitate the style of the adjacent text,[4] and a guild of reciters may maintain a very similar "house style" over centuries.

High readings, from **0·75** to **0·99**, are less compatible with a hypothesis of same author or same source. And when the D numbers reach or exceed **1·00**, we have the **Extreme** level, which forbids any such hypothesis: there is genuine discontinuity.

Some patterns recur often enough to have names. When the first and third of three passages are stylistically close, but the middle one is not, we have a **gap**, suggesting that the middle piece is intrusive. When several passages have the same "lookback," we may have a **horizon**, implying that they are looking back to the same earlier point – which may not be the beginning of that text as we now have it.[5]

Any proposed test must be calibrated against known examples, where there is a consensus, or such a persuasive argument that we are reasonably sure of the answer. If the test seems to be seeing what competent readers have been seeing, then any new information it may provide will be welcome.

Whole Texts

1. **The Pentateuch**. These five texts are conventionally ascribed to Moses, but scholarly opinion entirely rejects Mosaic authorship. The BIRD results are as follows (word counts, to the nearest thousand, are given in the second row):

	Gen	*Exod*	*Lev*	*Num*	*Deut*
	28k	23k	17k	23k	20k
Gen	~	**1·05**	**1·87**	**1·15**	**1·94**
Exod	**1·05**	~	**1·03**	**1·00**	**1·76**
Lev	**1·87**	**1·03**	~	**1·21**	**1·51**
Num	**1·15**	**1·00**	**1·21**	~	**1·59**
Deut	**1·94**	**1·76**	**1·51**	**1·59**	~

. . . and every single two-test result is at or above the Extreme level. The scholarly verdict is handsomely confirmed. Not only did the same person not write these texts, *no one person wrote any two of them*.

This will surprise few experienced readers, who are accustomed to the idea that no one person wrote any *one* of them. We are just getting started.

---

[4]Or we may have a *generic* similarity, as between the genealogies in Genesis 5, 10, 11, 36.
[5]Thus Gen 45-47 apparently invoke the early Jacob narratives of Gen 28-32.

2. **Premonarchic History**. There next follows the premonarchic history of Israel. In some Bibles, this comprises three texts: Joshua, Judges, Ruth:

	Joshua	Judges	Ruth
	15287	13879	1788
Joshua	~	0·73	**1·07**
Judges	0·73	~	**1·08**
Ruth	**1·07**	**1·08**	~

The first two are on the high side of Normal; in the same stylistic ballpark.[6] But Ruth is prohibitively different from both. This agrees with its character: Ruth is not history; it is a response to the Ezra/Nehemiah prohibition of foreign wives. It makes its point by reminding everyone that David himself was descended from a foreign woman.

3. **Ruth**. How then does Ruth compare with Ezra and Nehemiah (omitting the Aramaic portions of the latter two)? The answer is. . .

	Ruth	Ezra	Neh
	1788	3693	7711
Ruth	~	0·88	**1·01**
Ezra	0·88	~	0·59
Neh	**1·01**	0·59	~

. . . that Ezra and Nehemiah are reasonably close (the near side of Normal), and might belong to the same school. But Ruth is distant from both, especially from Nehemiah. Ruth undoubtedly had one or both of them"in mind," but not in any very friendly way.

4. **Monarchic History** proper is found in Samuel and Kings, which were earlier regarded as the Four Books of Kings. 1 and 2 Kings are close enough to count as chapters in a continuous narrative. Not so 1 and 2 Samuel, which are close enough to be chapters in the same story, but not parts of a single consecutive story:

	1 Sam	*2 Sam*	*1 Kgs*	*2 Kgs*
	18641	15413	18398	17101
1 Sam	~	0·70	0·92	**1·09**
2 Sam	0·70	~	0·87	0·86
1 Kgs	0·92	0·87	~	0·48
2 Kgs	**1·09**	0·86	0·48	~

The only modest similarity of 1 and 2 Samuel may be due to their containing various proportions of popular material,[7] not to mention the role of Samuel himself, who represents the antimonarchic strain in Judges, whereas Kings asks only one question: Will this king condone worship of the Baals, or not?

---

[6]Those to wish to ask the Hexateuch question will ask, Does Joshua pick up smoothly from Deuteronomy, as that theory might suggest? The answer is No (D = 1.30).

[7]Such as the twin narratives of the Goliath story; see Brooks **Goliath**.

5. **Revisionist History**. Chronicles is a Temple-centered remake of Kings, with some distinctive new material. The changes from Kings are inconsistent in kind and amount, from chapter to chapter. It was thus unlikely that 1 and 2 Chronicles would be stylistically similar to each other. And they are not:

	1 Kgs	2 Kgs	1 Chr	2 Chr
	18398	17101	15317	19426
N001	~	0·45	**1·52**	**1·17**
N002	0·45	~	**1·57**	**1·29**
N003	**1·52**	**1·57**	~	0·93
N004	**1·17**	**1·29**	0·93	~

The D value for the two is 0·93, just short of Extreme. With Kings, despite some identical material, all Chronicles relations *are* Extreme. Chronicles is indeed a remake.

6. **Isaiah**. To Duhm (1892) is due the suggestion that Isaiah is not one book, but three (Isa 1-39, 40-55, and 56-66). A first reaction might be to ask, Do those three parts differ from each other in a way that would support the Duhm hypothesis?

	1 Isa	*2 Isa*	*3 Isa*
	13271	5638	3608
1 Isa	~	0·65	0·50
2 Isa	0·65	~	0·50
3 Isa	0·50	0·50	~

Yes, they do. Second (Exilic?) and Third (post-Exilic?) are close to each other, and Third Isaiah also "looks back" to First Isaiah, as though being mindful of its origins. If Isaiah is the work of a continuous school, this is one way it might look.

But there is more: each of the three parts has its own structure. In First Isaiah, the prophet is called only in Isa 6; what about the stuff before that?. Isa 36-39 parallels the account of Hezekiah in 2 Kings 18-20; who is reading whom? Books are written on the problems of Isa 24-27, which has connections not just with 2 Kings, but with many texts, some (like Dan) obviously late. Let us then consider the Three Isaiahs, but with Isa 1:24-27 omitted (giving 1 Isar; "r" for "residue"). With that adjustment, we get:

	1 Isar	*2 Isa*	*3 Isa*
	12194	5638	3608
1 Isar	~	0·57	0·44
2 Isa	0·57	~	0·50
3 Isa	0·44	0·50	~

And *all three sections* come closer together. We have just established that Isa 24-27, whatever its internal logic, is a late addition to Isaiah (and not an intrusion of, say, Third Isaiah into an earlier First Isaiah).

Such are the benefits of working on passages within texts, as well as whole texts. To such problems, the rest of this study will be devoted.

Passages Within Texts

7. **The Creation Stories**, Gen 1 (creation in seven days) and Gen 2-3 (Eden), are manifestly distinct. Gen 4 carries the story of Adam and Eve to that of Cain and Abel. We expect Gen 1 to be distinct from the three others, and Gen 2 and 3 to be especially close to each other. And that is exactly what we find:

Gen	*1*	*2*	*3*	*4*
	662	491	494	450
1	~	0·62	**1·00**	0·95
2	0·62	~	0·49	0·71
3	**1·00**	0·49	~	0·56
4	0·95	0·71	0·56	~

The core chapters of the Eden story (Gen 2-3, **0·49**) are closely similar, implying continuous narrative; Gen 4 (at **0·56**, Normal) is intelligible as a related but separate episode in the same sequence. But Gen 1 is not close to Gen 3 (D = 0·62), and has an Extreme relation (D = **1.00**) to Gen 3. It is thus confirmed that the two stories are distinct. Gen 1 is a creation story of a widespread type, while Gen 2-3, the Eden story, though it too has parallels elsewhere, is more an aetiology of humankind.

8. **Gen 5 and 6**. It is customary to take Gen 1-11 as the first of three sequences in Genesis, telling of the origin of the Hebrew people. A problem with that lineage is that when Abel is killed and Cain is banished, there is no one to carry on the lineage.

That problem is solved in Gen 5, with the birth of Seth. With Seth the Adamic lineage resumes, and continues with the story of Noah, Gen 6-9. Can BIRD tell us anything about the smoothness of that line of descent? Here is the table for Gen 4-7:

	4	*5*	*6*	*7*
	450	481	429	497
4	~	0·59	0·73	**1·08**
5	0·59	~	**1·14**	**1·23**
6	0·73	**1·14**	~	0·63
7	**1·08**	**1·23**	0·63	~

Gen 4 -5 are plausible consecutive episodes (D = 0·59); so are Gen 6-7 (D = 0·63). But the transition from Gen 5 to Gen 6 (D = **1.14**) marks a stylistic break. Adding the previous result, we have three stylistically distinct sections in Gen 1-11. (1) Gen 1, with its Near Eastern parallels; (2) Gen 2-5, the Eden story, with other Near eastern parallels, continued with Seth. (3) Gen 5-7, Noah's Flood. That story is universal,[8] with an especially close version in the Babylonian Enuma Elish.[9]

9. **Joseph**. The romantic story of Joseph (Gen 37-50) should also be sharply distinguished from the genealogies preceding it.

---

[8]For the complication that the Genesis Flood is actually *two* stories, see Brooks **Noah**.
[9]Speiser **Genesis** 9-11.

But the test seems to disappoint us:

Gen	35	36	37	38
	497	612	673	568
35	~	0·63	0·44	0·69
36	0·63	~	0·40	0·84
37	0·44	0·40	~	0·65
38	0·69	0·84	0·65	~

We expect contrast at the beginning of the Joseph story. Instead, we get continuity. What has happened? Most probably, whoever patched the Joseph romance onto the preceding Genesis story was concerned to link Genesis with Exodus as part of a new concept, the Hexateuch, which joined the Pentateuch and the Conquest narrative (Joshua) as a single long narrative. It may then have been precisely to conceal the abruptness of the stylistic transition at Gen 36 and 37 that the writer, with his eye firmly fixed on Gen 36, has fashioned his Gen 37 *so as to match it.*

The moral of this is that, as with the "lookback," a writer with his mind on some passage, whether intentionally or inadvertently, can produce a closely similar passage. Style does not just occur. It can sometimes be manipulated (more or less successfully) for a desired affect, as part of the purpose of any text: to persuade its readers.

10. **Leviticus 2** is seen by some as integral to Lev 1-3;[10] by others as intrusive.[11] BIRD, agreeing with the latter, sees Lev 1 and 3 as very close, with a gap between::

Lev	1	2	3
	362	276	354
1	~	0·66	0·37
2	0·66	~	0·92
3	0·36	0·92	~

Lev 1 and 3, on meat sacrifices, are then not yet assimilated to the cereal offerings of Canaan, which are already mentioned in Lev 4. Including that chapter, we get:

Lev	1	2	3	4
	362	276	354	774
1	~	0·66	0·37	0·36
2	0·66	~	0·92	0·97
3	0·37	0·92	~	0·53
4	0·36	0·97	0·53	~

We then have Lev 1 and 3 on meat sacrifices, and Lev 4 on meat sacrifices *and* cereal offerings, stylistically near to Lev 3, and "looking back" (as does Lev 3) to Lev 1.

---

[10]Milgrom (2004) 25, ". . . would tend to support the view that the two are related."

[11]Noth (1962) 26, "Chapter 2 has been fitted in between chs 1 and 3, which belong closely together and correspond exactly." This is the picture which BIRD is also seeing.

What compositional scenario might underlie this? One possibility is an original treatise on meat sacrifices, Lev 1 and 3, then a later, stylistically compatible Lev 4, which includes cereal offerings, and finally, tidying up the treatise, the added Lev 2, inserted at the point where cereal offerings would logically have been first mentioned.

11. **Leviticus 15**. The "gap" pattern presented by Lev 1-3, with the ends closely similar and the middle item less so, occurs once more in Leviticus, at Lev 14-16:

Lev	14	15	16
	1265	649	812
14	~	0·87	0·32
15	0·87	~	0·76
16	0·32	0·76	~

Lev 15 is about impurity resulting from bodily discharges. Its originality has been challenged.[12] The test results suggest a work extending itself by additions at the end, but with some afterthought inclusions, in the interest of overall completeness.

12. **Balaam's Story**, at Num 22-24, draws on and incorporates a known person, the seer Balaam. That story exists in two versions: a popular one, in which an Angel of the Lord confronts Balaam, and but for the keener vision of his talking ass, would have killed him (part of Num 22), and a dignified one, where Yahweh's approval is given before Balaam sets out (Num 23-24). Does BIRD distinguish these versions?

Num	22	23	24
	866	516	411
22	~	0·90	0·81
23	0·90	~	0·41
24	0·81	0·41	~

Yes, it does. There is a contrast between a popular and a learned account of Balaam.[13] The popular account has probably been somewhat rewritten for inclusion in Numbers, but apparently not so much as to obliterate its different literary origins.

13. **Numbers** 36:13 ends Numbers thus: "These are the commandments and the ordinances which Yahweh commanded by Moses unto the children of Israel in the plains of Moab by the Jordan at Jericho." Before that is inserted a ruling: a challenge to a previous case (recorded in Num 27:1-12), on the rights of daughters (in this case, the daughters of Zelophehad) to inherit.

It has been thought strange that these obviously related cases appear at different places in Numbers. Noth 257 calls **36:1-12** "an appendix."

---

[12]Noth 113 "it does not deal with a subject of special priestly professional knowledge." Milgrom 140, "Why are the two chapters that discuss genital discharges separated? If this portion of Leviticus were organized thematically, one would have expected chapter 15 to be joined to the other genital-discharge chapter, chapter 12."

[13]For Balaam and the Deir Alla inscription, see van Kooten et al **Prestige**. The two parts of the Balaam story are further discussed in Brooks **Balaam**.

We may ask, How stylistically similar are the summaries of these two related cases? With "27x" standing for "an extract from Num 27"), and so on, we have:

Num	27x	36x
	211	266
27x	~	0·57
36x	0·57	~

The D value, 0·57, is at the near end of Normal. As far as this result goes, these two cases might plausibly have been taken from the files of the same court of record.

That may indeed be the likeliest scenario. And since the second was not appended to the first, as everyone naturally expects, it may instead have been added simply when it became known to the writer of that part of Numbers. Then the distance between the two in Numbers (the distance from chapter 27 to 36) may in some way correspond to the time between the two actual cases. How long might it have been, before someone would bring a countersuit on behalf of the clan interest in preserving clan land? Maybe not a year, but also not a generation. Since the daughters of Zelophehad are seemingly still unmarried, the best guess will be that it followed rather soon. However long it was, it might be a measure of the rate of extension of Numbers in that period. If it was four years, then the Numbers rate is one chapter every six months. That may well seem a reasonable, indeed, a leisurely, rate at which to write, or rewrite, the Numbers history of the Conquest, which is what occupies its author in Num 28-35.

14. **Second Zechariah**. It has long been seen[14] that Zechariah is actually two texts, one consisting of Zech 1-8 and the other of Zech 9-14. The stylistic difference is:

Zech	1-8	9-14
	2454	1938
1-8	~	1·00
0-14	1·00	~

And the common opinion is decisively confirmed.

But suppose we were unsure where the division occurred? We could then test every possible division, from Zech 1 vs 2-14, to Zech 1-13 vs 14, and choose the one that maximized the stylistic difference. If we did, the D numbers would be the following:

Zech 1 vs 2-14	0·71	Zech 1-5 vs 6-14	0·68	Zech 1-9 vs 10-14	0·86
Zech 1-2 vs 3-14	0·77	Zech 1-6 vs 7-14	0·78	Zech 1-10 vs 11-14	0·95
Zech 1-3 vs 4-14	0·78	Zech 1-7 vs 8-14	0·87	Zech 1-11 vs 12-14	0·79
Zech 1-4 vs 5-14	0·69	Zech 1-8 vs 9-14	**1·00**	Zech 1-12 vs 13-14	0·84
				Zech 1-13 vs 14	0·89

And the division specified by previous scholarship turns out to be also that of maximum stylistic difference.

If we had to locate that dividing point by the test alone, without reading anything, we would put it exactly where our predecessors did.

---

[14]Going back to Joseph Mede (1632); see Smith **Twelve** 2/440.

15. **Third Zechariah**. Scholarship has also identified a transition between Second (Zech 9-11) and Third Zechariah (Zech 12-14). The BIRD results are:

Zech	9	10	11	12	13	14
	306	226	337	321	214	534
9	~	0·34	**1·07**	0·83	0·72	0·95
10	0·34	~	0·98	0·78	0·64	**1·01**
11	**1·07**	0·98	~	**1·01**	0·56	0·82
12	0·83	0·78	**1·01**	~	0·80	0·93
13	0·72	0·64	0·56	0·80	~	0·72
14	0·95	**1·01**	0·82	0·93	0·72	~

Following down the diagonal, the only Extreme value we get is the D = **1·01** reading between Zech 11 and 12. Previous scholarly opinion is confirmed – though with a suggestion that the continuity of Zech 11 with Zech 10 could maybe use another look.

A Concluding Word of Caution

We naturally seek for "authors," but style as such does not reflect authorship; one author can write in many styles, from elegaic to jocular. What BIRD gives us is simply a measure of difference. If it is low, we have consecutive narrative, or the writer "has in mind" the other passage. If it is extreme, we have a different author or source. There are many ways in which two passages can be either similar or different. Being based on high-frequency words, BIRD can work with small samples, but as the size of the passages tested decreases, the test words become less operative, and the chance of false positives steadily increases. Nor are these stylistic results self-interpreting; interpretation must always remain the responsibility of the fallible investigator.

But whatever its limitations, BIRD is not influenced by any human expectations: *it simply tells us what it sees.* That negative is the test's strongest recommendation. That it sometimes agrees with what seems like sound previous opinion merely tells us that those previous scholars and BIRD are, after all, looking at the same thing.

Works Cited

E Bruce Brooks. Balaam. Alpha v3 (2022) 00-00

E Bruce Brooks. Goliath. Alpha v2 (2021) 53-54

E Bruce Brooks. Noah's Flood. Alpha v3 (2022) 00-00

E Bruce Brooks and A Taeko Brooks. Stylistic Difference in Chinese and Greek. Alpha v2 (2021) 15-30

Bernhard Duhm. Das Buch Jesaia. 1892; 2ed Vandenhoeck & Ruprecht 1902

J Scott Hibbard. Intertextuality in Isaiah 24-27. Mohr Siebeck 2006

George H van Kooten and Jacques van Ruiten. The Prestige of the Prophet Balaam. Brill 2008

Jacob Milgrom. Leviticus. Fortress 2004

Martin Noth. Leviticus.1962; tr Westminster 1966; rev 1977

George Adam Smith. The Book of The Twelve Prophets. 2ed Harper 1928

F A Speiser. Genesis [AB]. Doubleday 1964