Untangling the derivatives: points for clarification in the findings of the Shakespeare Clinic

Thomas Merriam

Abstract

The work of the Shakespeare Clinic of Claremont McKenna College, led by Ward E.Y. Elliott and Robert J. Valenza, is recognized for its pioneering computer analysis of many early modern texts to determine whether William Shakespeare (1564–1616) wrote the works traditionally ascribed to him. The Clinic achieved its primary objective of eliminating all other known candidates and thus confirming that Shakespeare wrote them. Two general methods of analysis were applied to whole plays and variable-sized large texts: Discrete Composite Analysis and Continuous Composite Analysis.. The first uses univariate analysis to determine acceptance or rejection of forty-eight stylometric tests for each text. The second uses a multi-dimensional composite mean for Shakespeare derived from all forty-eight in order to determine acceptance or rejection for each text. This article notes the omission of Discrete Analysis to take into consideration statistical dependencies between the forty-eight tests, the partly arbitrary 'handfitting' of acceptance–rejection boundaries for each of the forty-eight tests, the failure to take into full account the factor of chronology, and the absence of discussion of the part played by prior probabilities as to existing beliefs concerning attribution. By this last point, I mean the role played by the existing traditional consensus as to Shakespeare attribution, prior to linguistic analysis. For Continuous Analysis, it is noted that the stated probabilities are not true probabilities as acknowledged, and that the resulting acceptance–rejection levels for them are calibrated in line with prior beliefs. Principal component analysis is shown to give improved results in dealing with co-authored Shakespeare plays, *Henry VIII*, *Timon of Athens*, and *Pericles*. This does not invalidate the overall aim of the Shakespeare Clinic.

Correspondence:
Thomas Merriam
35 Richmond Road,
Basingstoke RG21 5NX, UK.
E-mail:
merriam12484648@
hotmail.co.uk

According to Professor Ward Elliott and Robert Valenza, over the years, critics of their Shakespeare Clinic's findings are of two kinds, the favourable and the not so favourable. Those favourable consist of scholars who liked their conclusions and their methods 'such as Don Foster before 1986 and the Oxfordians before 1990', as well as 'most of the thirty-odd outside scholarly readers of our journal articles.' The not so favourable consist of those who did not like the conclusions and those who did not like their methods 'such as literature department numerophobes who think that crunching Shakespeare is as gauche and perverse as drinking from the finger bowl' (Elliott and Valenza, 2004, p. 362).

Such categorization into 'all for' or 'all against' implies that critical assessment of the Clinic's methods is consequent on prior judgment of its

Literary and Linguistic Computing, Vol. 24, No. 4, 2009.
doi:10.1093/llc/fqp026 Advance Access published on 20 July 2009

conclusions. I would suggest, on the contrary, that it is possible to welcome some of the Clinic's conclusions and not others and to approve of some of the Clinic's methods and not others. Relegating all criticism to the categories of pro and con fails to do justice to MacDonald Jackson, one 'of the [two] greatest masters of authorship studies' (Elliott and Valenza, p. 361), whose disagreement with Elliott and Valenza over the authorship of Hand D in *Sir Thomas More* is based on painstaking attention to the Clinic's methods (Jackson, 2006).

At the outset, it is important to remember that the founding purpose and original élan of the Shakespeare Clinic of Claremont McKenna College was to determine whether or not the author of Shakespeare's plays and poems was William Shakespeare, born and died in Stratford-upon-Avon, 1563–1616. Those who believe that he was the author are termed 'Stratfordians' by the 'anti-Stratfordians' who believe that 'Shakespeare' was a *nom-de-plume* for another author deemed as worthier of literary genius. Professor Elliott's father was an anti-Stratfordian, and he himself was similarly inclined when he inaugurated the Clinic's lengthy and arduous researches.

Although in the United Kingdom the cause of the anti-Stratfordians has come to be regarded as *passé* at best, in the USA it claims, according to Professor Elliott, adherents among some in the legal professions. It is to Ward Elliott's great credit that his researches led him to acknowledge insufficient evidence in favour of any of the possible anti-Stratfordian claimants, and thus change his mind.

Elliott and Valenza succeeded in demonstrating that no other claimant wrote a core canon of twenty-nine Shakespeare plays. This achievement can be overlooked by those for whom the Stratfordian issue is not of primary importance. The difficulty with the work of the Clinic lies not in its original purpose and accomplishment, but in auxiliary considerations, many which defy straightforward explanation. What follows is an appeal for further and deeper consideration.

An essential first step in comparing Shakespeare plays with non-Shakespeare plays is the screening of a core canon of texts which may be assumed as Shakespeare's alone. Elliott and Valenza began with thirty-eight plays. These were subsequently reduced by eliminating *1 Henry VI*, *Timon of Athens* and all of *Henry VIII*, while retaining 'Shakespeare's' parts of *Pericles* and *The Two Noble Kinsmen* (Elliott and Valenza, 1996, p. 211). Further pruning led to the establishment of a twenty-nine play 'clean, commonized baseline' with the elimination of *2 Henry VI*, *3 Henry VI*, *Henry V*, all of *Pericles* and all of *The Two Noble Kinsmen* (Elliott and Valenza, 2004, p. 399).

The Clinic's tests of authorship consist of seventeen 'Round One Tests' noted as 'new-tech, hyphenated words' (Elliott and Valenza, 2004, pp. 408–9), sixteen 'Round Two Tests' consisting of 'contractions, metric fillers, selected words, and phrases, per 20,000 words' (Elliott and Valenza, 2004, p. 412), and fifteen 'Round Three Tests' consisting of 'prefixes, suffixes, intensifiers, and adversions, per 20,000 words' (Elliott and Valenza, 2004, p. 416). Using these forty-eight stylistic criteria, Elliott and Valenza provided two statistical analyses, the first of which, Discrete Composite Analysis, uses univariate statistics to establish a null hypothesis of each of the forty-eight tests. The null hypothesis requires the data for each individual test to fall (or not to fall) within parameters that conform to those set by the twenty-nine play 'clean, commonized baseline', that is, the Shakespeare profile. The verdict for each of the forty-eight tests is an acceptance or a rejection. An acceptance does not evidence Shakespearian authorship; a rejection indicates a non-Shakespearean element in the forty-eight element profile. The second method, Continuous Composite Analysis, is one which is described as follows:

> This method (1) aggregates every Shakespeare's mean on every test into a multi-dimensional composite mean; (2) then measures a given text block's distance, in standard deviations, from Shakespeare's mean on every test; and (3) then aggregates the 'Shakespeare distance' on every test into a 'Continuous Composite Error' ('CCE') score (Elliott and Valenza, 2004, p. 350).

The results of both analyses are expressed in units of probability, but they are not probabilities as

commonly understood. Table entries such as the 120 occurrences of '<1.000E–15' in 'Oxford by the Numbers' are expressed in statistical tables as simply 'less than one-half of 1%', or < 0.5%.

> It is important to stress again that composite probability scores, whether from Discrete or Continuous analysis, are not indicators of the absolute, actual probability that Shakespeare wrote the block in question. Composite probability scores are markers from which composite Shakespeare ranges are derived. The scores permit *comparison* of the block in question, not with the closest theoretically imaginable Shakespeare match, but with an actual Shakespeare block at the edge of his range (Elliott and Valenza, 2004, p. 351).

Elliott and Valenza present an overview of the strengths and weaknesses of both analytical methods. Discrete Composite Analysis has the virtue of being 'easy to understand, compute, and present' (Elliott and Valenza, 2004, p. 355). The Discrete Composite probabilities for the twenty-nine plays which constitute the Shakespeare baseline, or core canon, range between 1.0 and 0.2316, well within the accepted limits of the null hypothesis. But that in itself 'only gives one the crudest of notions of the odds of Shakespeare authorship . . . ' (Elliott and Valenza, 2004, p. 355).

For the Elliott-preferred Discrete Composite Analysis, each of the forty-eight tests for the Shakespeare baseline is evaluated according to its number of acceptances and rejections within its respective twin set of high and low parameters. How are these parameters determined? The preferred method is to determine the means and standard deviations, once given the set of twenty-nine core canon plays and then verify which values fall within (or without) two standard deviations of the mean, plus or minus. Professor Elliott has used instead an *ad hoc* approach or so-called 'handfitting', whereby parameters are set generally according to the maxima and minima range of the twenty-nine plays. An element of circularity enters into this procedure, as suggested by the anomaly of *Much Ado About Nothing* with a Discrete Composite probability of 0.6018 (one rejection) based on 'custom fitting',[1] or 0.0045 (five rejections) based on two standard deviations from the mean. As *Much Ado About Nothing* is accorded a strong prior probability as an all-Shakespeare play, its Discrete Composite probability can only fall within the acceptable limits of the null hypothesis.

Additional difficulties, using the preferred method, are apparent with *Richard II* (0.013), *The Merry Wives of Windsor* (0.0045), *Hamlet* (0.013), and *Coriolanus* (0.013). The establishment of the twenty-nine play 'clean, commonized baseline' for Shakespeare is therefore as much a product of prior probabilities owing to received opinion of Shakespeare scholars, as it is based on the numerical methods of analysis used by the Clinic. *Richard II* (0.033822), *The Merry Wives of Windsor* (0.012563), *Much Ado About Nothing* (0.009988), and *The Tempest* (0.0036895) also present difficulties for inclusion, if one interprets the results of Ccontinuous Composite probability as conventional probabilities. The relatively early *Richard II* and the late *The Tempest* contain unaccounted-for authorship features, possibly collaborative inputs; *The Merry Wives of Windsor* and *Much Ado About Nothing* are unusual in having 87% and 72% prose, respectively. The discrepancy between the prior probabilities of received scholarly consensus and the findings of linguistic analysis, in principle call for further investigation from both sides.

An equally serious reservation about Discrete Composite Analysis concerns its use of univariate statistics in making an overall tally for the forty-eight tests. Individual tests are never 100% statistically independent as Elliott and Valenza tacitly assume them to be with discrete analysis. The correctly adjusted weight of contributing evidence is not the same for each of the forty-eight tests. Two tests may parallel each other to such an extent that their evidence constitutes less than that of two independent tests. Yet the Clinic's tally of rejections treats each rejection as equal. The way to obviate this unavoidable drawback of univariate statistics is to adopt multivariate statistics (principal component analysis).[2] Multivariate analysis generally confirms the homogeneity of the twenty-nine play baseline or core canon, and thus reinforces and

confirms the prior probability derived from conventional scholarship.

An omission of the Shakespeare Clinic's analysis of its carefully assembled data is a thorough treatment of the influence of time of composition upon that data. The authors of 'Oxford by the Numbers' acknowledge the inability of Continuous Composite Analysis to take the factor of time into consideration.

> ... Continuous does not account for time periods for traits like line endings, where Shakespeare's style changed over the years. Discrete analysis distinguishes between early and late profiles. Continuous does not (Elliott and Valenza, 2004, p. 350).

In fact, neither of the two methods comprehensively traces the effect of composition times on the results. For Continuous Composite Analysis, date of composition is not taken into consideration. For Discrete Composite Analysis, it is treated on an *ad hoc* basis. Elliott and Valenza have specified early and late (and sometimes, early, middle, and late) parameters for eight of the forty-eight tests which are obviously influenced by date of composition. These tests are feminine endings and open lines in Round One, 'on't', 'th'', 'i'th'', and ''d-'ld' in Round Two, and 'very' and 'most' + modifier in Round Three. Only the first two are 'traits like line endings'.

In all, there are seventeen of the forty-eight tests in 'Oxford by the Numbers' which have values for the core baseline which correlate significantly with the given dates. They are rare words, feminine endings, open lines, 'the 2lws', and BoB 7 in Round One; total 1, 'on't', 'th'', 'i'th'', ''ll', ''d-'ld', and ''tis' in Round Two; 'whereas | whenas', '-able', 'most' + modifier, 'see', and 'hark | listen' in Round Three. All seventeen that could make use of the subtabulation accorded the eight tests recognized as being dominantly time dependent. The dates of composition of every play tested by the Clinic might then be checked before evaluating the results of the seventeen tests. To do this would render the already complex tabulation more complex, thus subverting the claim that the process of Discrete Composite Analysis is 'simple and the results are easy to understand, compute, and present.'[3] This counsel of perfection may be superfluous, but further clarification of the matter of time dependency would help an improved evaluation of the method.

At this point, visual comparisons are apposite. Relevant data are tabulated in Appendix 1 for Figs 1–3. Figure 1 is a diagram which plots the Discrete Composite probabilities for the twenty-nine play Shakespeare baseline against the dates of composition provided by the authors of 'Oxford by the Numbers'. There are three tiers of probability, 0.2316, 0.6018, and 1, while the dates of composition number seventeen. A regression line is drawn that indicates a low level of correlation between dates of composition and the Discrete Composite probabilities. The coefficient of 0.184 for $n = 29$ is not significant of a correlation between the test results and date of composition. That a significant correlation actually exists for over one-third of the forty-eight tests is demonstrable. Discrete Composite Analysis fails adequately to reflect this far from a negligible factor.

Figure 2 plots the Continuous Ccomposite probabilities for the same plays against the same given dates of composition. The correlation coefficient of 0.229 for $n = 29$ is not statistically significant, although it is unexpectedly greater than for Discrete analysis, despite the disregard of chronology in Continuous Analysis. It still fails to reveal the importance of change over time in illustrating the 'clean, commonized baseline' of the twenty-nine plays.

Principal component analysis, which takes into consideration interdata dependence and 'telescoping' of all forty-eight tests for the twenty-nine plays, generates a first principal component with 17.7% of the total variation in the correlation matrix. As the values for the first principal component (PCA) are approximately normally distributed, they are transformed into corresponding probabilities which are plotted against the given dates of composition shown in Fig. 3. The correlation coefficient is 0.882 for $n = 29$, significant of a high degree of correlation between probabilities and dates. The Clinic's imperfect handling of time change creates an anomaly seen in the case of *Henry VIII* below.

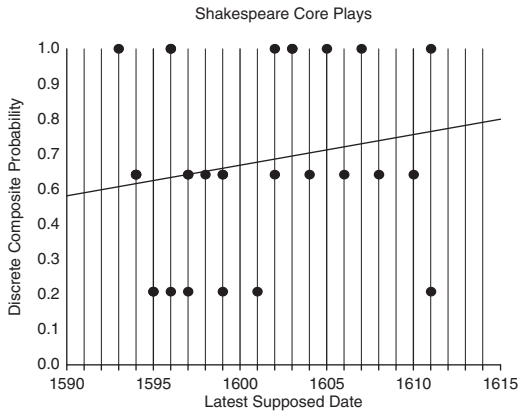

Fig. 1 Chronology as correlated with Discrete Composite probability for 29-play Shakespeare core canon

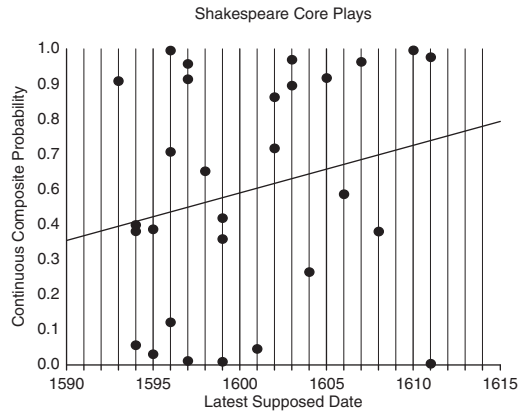

Fig. 2 Chronology as correlated with Continuous Composite probability for 29-play Shakespeare core canon

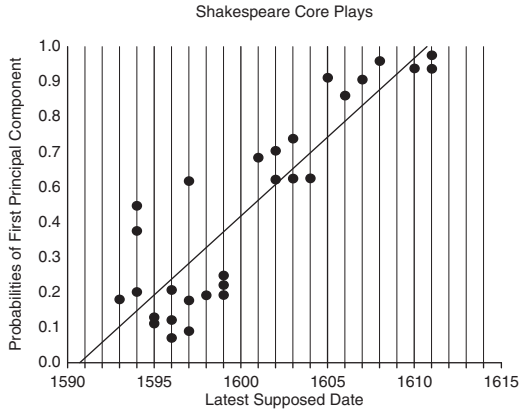

Fig. 3 Chronology as correlated with First Principal Component (PCA) for 29-play Shakespeare core canon

Elliott and Valenza divide *Henry VIII* into three sections entitled H8(Sh), H8(Jt for Joint), and H8(Fl). According to the electronic files shared with the author by Professor Elliott, H8(Sh) consists of 1.1 entire, 1.2 entire, 2.3 entire, 2.4 entire, 3.2.1–3.2.203, 3.2.350–3.2.439, 4.2.1–4.2.82, 4.2.108–4.2.173, 5.1 entire, and 5.2.182–5.2.215. H8(Jt) consists of 1.4.65–1.4.107, 2.1 entire, 2.2 entire, 4.1 entire. H8(Fl) consists of 1.3 entire, 1.4.1–1.4.65, 3.1 entire, 3.2.204–3.2.349, 4.2.83–4.2.108, 5.2.1–5.2.182, 5.3 entire, 5.4 entire, and epilogue (three Arden line numbering). This somewhat complicated division of the play is not made explicit in the authors' journal reports. All sections of *Henry VIII* are assigned to the category 'Dubitanda and set-asides', implying (correctly) that H8(Jt) and H8(Sh) are of mixed authorship. It will be noted that editors of the play have universally accepted Shakespeare as author of 1.1, 1.2, 2.3, 2.4, 3.2.1–3.2.203, and 5.1, which together constitute 81% of H8(Sh).

In a previous study, I determined that the matter included in H8(Jt) is 73% by Fletcher and the matter included in H8(Fl) is 97% by Fletcher. Seventy-nine percent of the material in H8(Sh) is by Shakespeare (Merriam, 2005). Elliott and Valenza's Discrete and Continuous Composite Analyses award Shakespeare authorship to none of the three divisions of *Henry VIII*. The twenty-nine play Shakespeare baseline is displayed in Figs 4–7 with open circles, while Fletcher plays are shown in black. Dates for the Fletcher plays are corrected in accordance with Gordon McMullan's entry for John Fletcher in *The Dictionary of National Biography*. Discrete Composite Analysis consigns H8(Sh), along with H8(Jt) and H8(Fl), to a level consistently below the *ad hoc* confidence level of 0.2316 established by the twenty-nine play baseline. Appendix 2 gives the particulars for Figs 4–7.

Continuous Composite Analysis also makes little distinction between the all-Fletcher H8(Jt) and

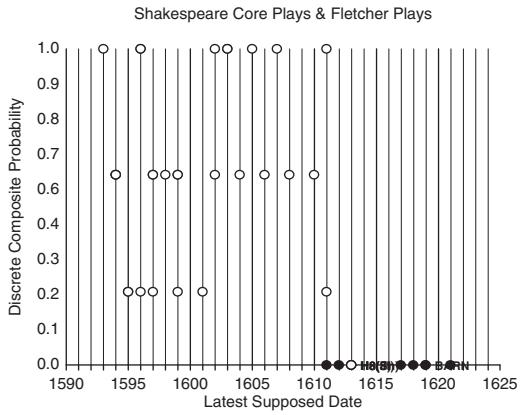

Fig. 4 Discrete Composite probabilities for Shakespeare, Fletcher plays and three parts of *Henry VIII* in chronological order

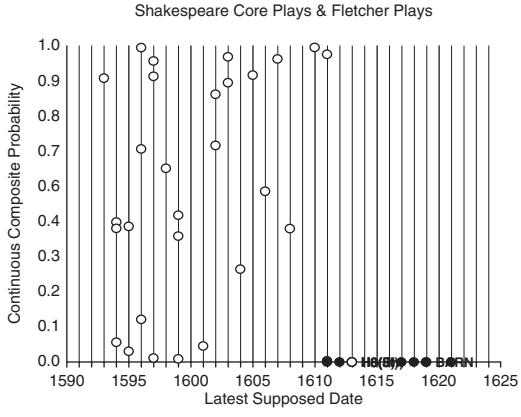

Fig. 5 Continuous Composite probabilities for Shakespeare, Fletcher plays and three parts of *Henry VIII* in chronological order

H8(Fl) on the one hand, and on the other, the mainly Shakespeare H8(Sh) shown with a single white circle at the base of the diagram. All parts of *Henry VIII* are positioned well below the 0.0036895 *ad hoc* level of acceptance permitted for Shakespeare authorship. It is misleading to say with regard to H8(Sh) that 'We believe that there are no glaring clashes between our findings and what is suggested by generally accepted documentary evidence' (Elliott and Valenza, 2004, p. 360). The testimony of Heminges and Condell as to Shakespeare's

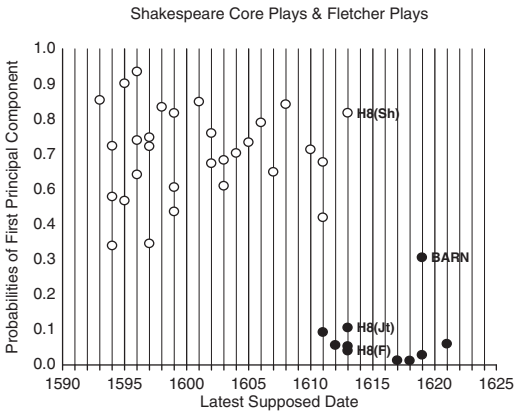

Fig. 6 First Principal Component probabilities for Shakespeare, Fletcher plays and three parts of *Henry VIII* in chronological order

authorship of *Henry VIII* comprises, at least in the parts universally accepted as Shakespeare's, 'generally accepted documentary evidence'.

PCA, in contrast, shows in Fig. 6 a clear divergence between H8(Jt) and H8(Fl) on the one hand, and H8(Sh) on the other. H8(Jt) has a probability nearer to that of the Shakespeare plays than the Fletcher plays, barring one collaborative play. Fletcher and Massinger's *Sir John Barnavelt* (1619) is labeled 'BARN' to draw attention to its separation from the other black-circled plays written solely by Fletcher. All the Fletcher plays are linearly separated from the Shakespeare baseline.

Most revealing is a plot of the second principal component in Fig. 7. Here, the correlation with time sequence for the Shakespeare plays is apparent; it involves a sequence which includes H8(Sh) but excludes H8(Jt) and H8(Fl) as well as the eight Fletcher plays. H8(Sh) fits appropriately at the base of the imagined regression line which connects the Shakespeare plays. H8(Jt) and H8(Fl) belong to a different associated cluster of plots. Fig. 7 illustrates the ability of PCA to encompass both authorship discrimination and chronological sequence.

What is the cause of the discrepancy between the Clinic's findings and those of PCA? H8(Sh) was plausibly rejected because of its positioning later than 1611, the last date within the Shakespeare baseline.

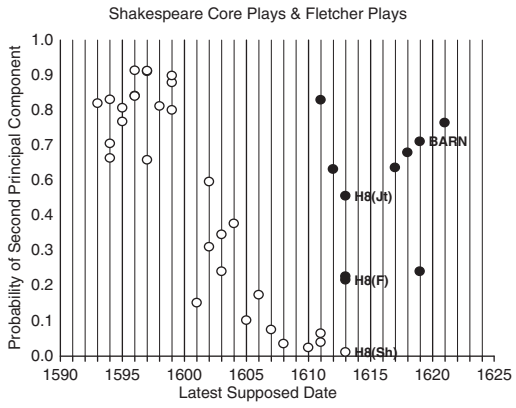

Fig. 7 Second Principal Component probabilities for Shakespeare, Fletcher plays and three parts of *Henry VIII* in chronological order

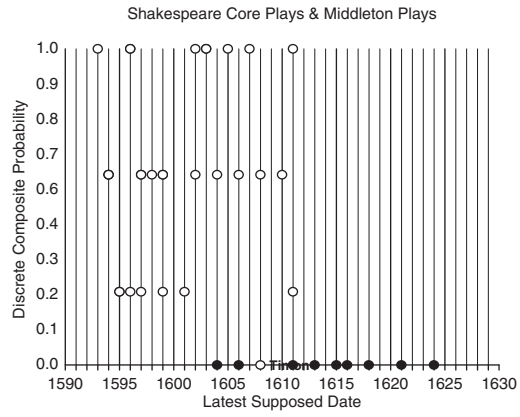

Fig. 8 Discrete Composite probabilities for Shakespeare, Middleton plays and *Timon of Athens* in chronological order

It must be said that had Elliott and Valenza strictly followed the principles of their authorship inquiry, *Henry VIII*, either whole or in part, would have been assigned to the category Apocrypha and not 'Dubitanda and set-asides'. H8(Sh)'s two given probabilities, 2.768E–07 and 5.234E–07 are comparable to those of the 'Claimant Play Discrimination' of *Edward II* with 1.337E–08 and 1.278E–07. Only because of prior probabilities based on literary criteria was *Henry VIII* included in the category of 'Dubitanda and set-aside', rather than removed from a Shakespeare association.

In the case of *Timon of Athens*, its rejections of 0.000000000004355 by Discrete Composite probability and 0.000000000000001 by Continuous Composite probability, the same probability as that of Thomas Dekker's *The Whore of Babylon*, should exclude it from the Dubitanda. The play is held to be a collaboration by Shakespeare and Thomas Middleton. The editor of the Oxford Middleton, Gary Taylor states:

> Middleton collaborated with the older playwright, writing about a third of *Timon of Athens*, including the bitterly comic central sequence where Timon's creditors turn their backs on him. Usually the most successful scenes of the play in performance, these apply classical tragedy techniques and materials developed in Middleton's recent city comedies (Taylor, 2008).

Timon (17,704 words) with two-thirds of its matter by Shakespeare is comparable to H8(Sh) (11,973 words) with four fifths by Shakespeare. In both cases, their test results should indicate major Shakespearean authorship. Using the methods of the Shakespeare Clinic, they do not.

Elliott and Valenza's analyses reject *Timon* as they do *Henry VIII* (Shakespeare). Figures 8 and 9 show their results for *Timon* and the eight listed Middleton plays (Elliott and Valenza, 2004, p. 400). Middleton plays are presented in black and Shakespeare baseline plays in white.

These results clash with 'generally accepted documentary evidence', based again on Heminges and Condell's inclusion of *Timon* in the 1623 Shakespeare First Folio. The Clinic's findings for H8(Sh) and *Timon* constitute false negatives, belonging to the type II or β error which Elliott and Valenza claim for the most part to avoid in their study of authorship.

> . . . our distinguishing stock-in-trade has been 'silver-bullet' evidence that *tends to* disprove common authorship by showing differences, rather than 'smoking gun' positive evidence used by most other analysts to prove common authorship with similarities such as

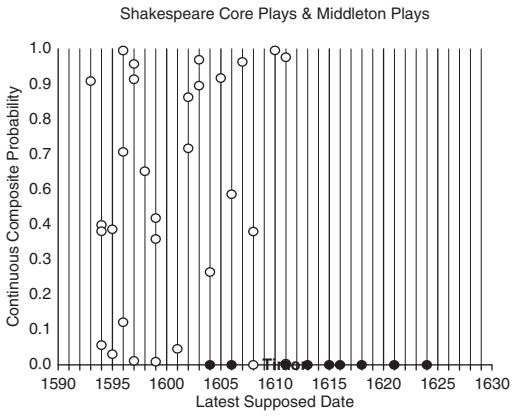

Fig. 9 Continuous Composite probabilities for Shakespeare, Middleton plays and *Timon of Athens* in chronological order

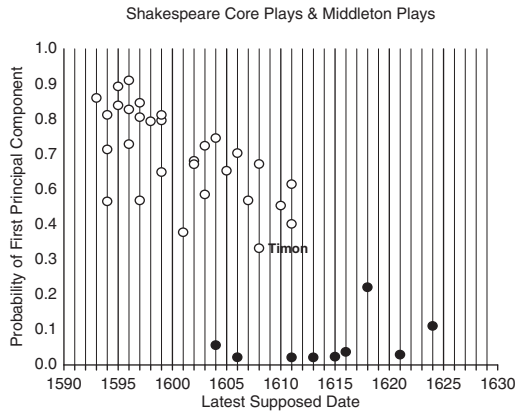

Fig. 10 First Principal Component probabilities for Shakespeare, Middleton plays and *Timon of Athens* in chronological order

> 'borrowings' or 'echoes' (italics mine) (Elliott and Valenza, 2004, p. 337).

PCA, on the other hand, presents a picture in Fig. 10 which harmonizes with the prior probabilities of literary criteria. Not only is the factor of chronological drift in the Shakespeare baseline plainly present, *Timon* is shown to be the Shakespeare play that is nearest to the cluster of black circled Middleton plays.

Included among the 'Dubitanda and set-asides' in 'Oxford by the Numbers' is *Pericles 3–5*, universally held to be by Shakespeare alone. Figure 11 below indicates the results of Discrete Composite Analysis.

Pericles 3–5 with a discrete probability of 0.06309 falls below the permitted confidence level of 0.2316 for the Shakespeare baseline, again type II error or false negative. On the other hand, it passes inside the Shakespeare threshold with Continuous Composite Analysis, having a probability of 0.006864 just within the lower acceptable confidence level of 0.003689.

PCA shows *Pericles 3–5* in Fig. 13 to be comfortably within the Shakespeare envelope and situated approximately on the diagonal which correlates its probability with date of composition.

Further examples could be provided for *Two Noble Kinsmen* (Sh's part), *Henry VI, Part 2*, and

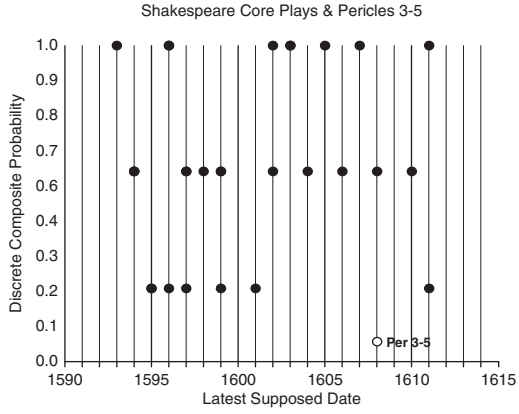

Fig. 11 Discrete Composite probabilities for 29-play Shakespeare core canon and *Pericles* 3-5 in chronological order

Titus Andronicus—late stratum, all listed among texts of the 'Dubitanda and set-asides'. Principal component analysis in these instances, as with those above, gives a more nuanced picture than Discrete and Continuous Composite Analyses. This is because Discrete and Continuous Analyses were designed to lead to a bivalent decision 'as to one of only two possible results, that is the acceptance or rejection of a hypothesis' (Howson and Urlach, 1989). They are unsuited to the gradations of Shakespearian co-authorship which are arguably

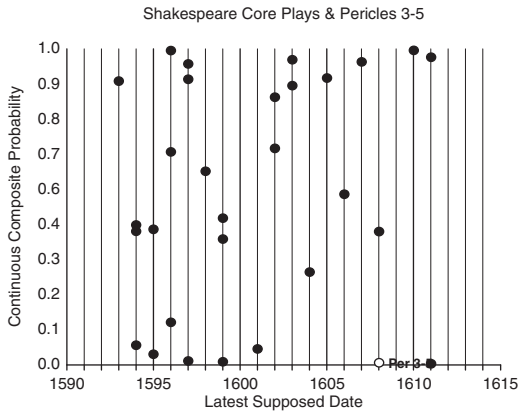

Fig. 12 Continuous Composite probabilities for 29-play Shakespeare core canon and *Pericles* 3-5 in chronological order

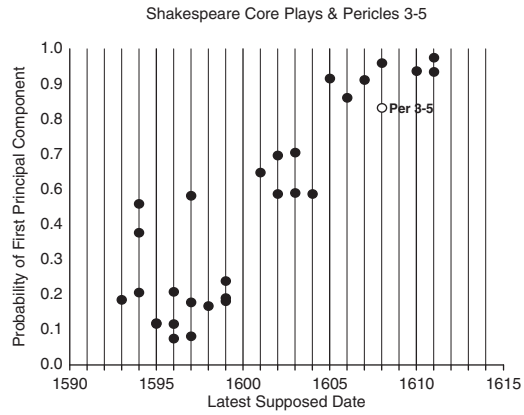

Fig. 13 First Principal Component probabilities for 29-play Shakespeare core canon and *Pericles* 3-5 in chronological order

present in varying degrees in the canonical Shakespeare history plays, along with *Titus Andronicus*, *Edward III*, *Sir Thomas More*, *Timon of Athens*, *Pericles*, *Macbeth*, and *Two Noble Kinsmen*. The originating thrust of the Shakespeare Clinic's work was directed, as previously stated, toward testing the hypothesis that the solo works of Shakespeare were written by a non-Shakespeare Claimant—yes or no.

> . . . the null hypothesis is never proved or established, but is possibly disproved in the course of experimentation. Every experiment may be said to exist only to give the facts a chance of disproving the null hypothesis (Fisher, 1947).

Fisher's formulation of classical statistics agrees with Popper's view that while theories cannot be proven by empirical evidence, they can sometimes be disproved by them. Elliott and Valenza's refutation of the hypothesis that any named Claimant wrote Shakespeare's core canon follows this template. The strategy used is one which regards positive empirical evidence of little or of no value; only negative evidence is utilized to refute the null hypothesis once formulated.[4]

In Professor Elliott's words:

> Our tests are more like comparing shoe sizes, blood typing, or eye color than comparing fingerprints. If our tests are defined and measured properly, they will show tons of false positives but no more than ounces or pounds of false negatives. We believe that negative evidence normally outweighs positive evidence by many orders of magnitude. As noted, fitting a tiny slipper does not prove you are Cinderella nearly as conclusively as *not* fitting the tiny slipper proves you are *not* Cinderella. If you are a size four, you could just as well be a false positive Little Miss Muffet or Tiny Tim; but, if you are a size ten, your claim to be Cinderella is in trouble. The trouble is compounded, and the disproof stronger, for every additional profile you do not fit – hat size, height, eye color, or blood type – making it easy to eliminate a Cinderella claimant even if uncanny numbers of other measurable features – sex, ring size, hair color, inseam, resting pulse rate, cholesterol level, or blood pressure – seem to match exactly. Hence, our distinguishing stock-in-trade has been 'silver-bullet' evidence that *tends to* disprove common authorship by showing differences, rather than 'smoking gun' positive evidence used by most other analysts to prove common authorship with similarities such as 'borrowings' or 'echoes' (last italics mine) (Elliott and Valenza, 2004, p. 337).

There are several difficulties with this argument. First, the evidence of shoe size and blood type, if not eye color, is measurable in a deterministic sense. One's shoe size, as it were, is either four or five, or ten. One has blood type A, B, AB, or O. The measurements of the Shakespeare Clinic, on the other hand, are probabilistic. Their status as decisive positive or negative evidence is dependent on parameters which can never be certain, even using conventionally approved limits like two standard deviations, and indeed less so with the subjective use of 'handfitting' the parameters.

The second difficulty is the assumption that emphasis on negative evidence in refuting a hypothesis is unassailable in logic. This involves the insoluble problem of deduction from induction. Fisher was careful to include the word 'possibly' in his statement quoted above. Popper included the word 'sometimes', and Elliott and Valenza have included the collocation 'tends to'. These are, nevertheless, words that are easily forgotten in the course of an argument backed by Elliott's metaphorically rich rhetoric.

A third difficulty is the failure of the authors to clarify what is meant by a false negative. The term itself reflects the either-or nature of the investigation, core baseline compatible or not. The authors assume, without being explicit, that a negative verdict entails a rejection of the established 'clean, communized baseline' of twenty-nine plays. By setting the threshold of the Continuous Composite Analysis at the lowest probability observed for the given baseline, namely 0.003689, they guarantee the existence of true positives and the absence of false positives, as pre-defined solely in terms of the core baseline. The calibration of Valenza's Continuous Composite probabilities eliminates all other plays from the baseline besides *Henry VI, Part 2*, with a probability of 0.2724 and *Pericles 3–5* with a probability of 0.006864, both having the values above the threshold of 0.003689.

On the other hand, if a negative verdict is defined as the rejection of a text which is not all or part by Shakespeare according to implicit prior probabilities on literary grounds, as in the case of *Henry VI, Part 1, Henry VI, Part 3, Henry V, Henry VIII, Two Noble Kinsmen, and Titus Andronicus*, then the resulting calibration of the Continuous Composite probabilities creates false positives of all the Claimant and Apocryphal plays. Some of the existing Dubitanda negatives have Continuous Composite probabilities which are infinitesimal, namely *Henry VI, Part 1*, the 10,609 words of *Titus Andronicus, Timon of Athens* entire, and Hand D of *Sir Thomas More*, all with less than 1E–15.

Seventy-five plays tested have an undifferentiated probability of less than 1E–15. It is hard to believe that such an equally vast difference from Shakespeare can be generated by so many different works in the same language. Had the calibration of Valenza's analysis been set to ignore gradations of authorial resemblance simply in order to avoid false positives?

Finally, there is the definition of false negatives and true positives in terms of the statistical probabilities as 'indicators of the absolute, actual probability that Shakespeare wrote the block in question'. This would follow the tradition of classical statistics, but would fail to confirm the authorial configuration of texts in 'Oxford by the Numbers' by rejecting several of the baseline plays as false positives.

Much the same can be said for Discrete Composite probabilities. These are set with a boundary threshold of 0.2316, thus eliminating all plays other than *Henry V* with the apparent false positive of a probability of 0.2316. As the number of rejections for individual plays making up the core baseline is based on 'handfitting', and as the probability threshold is calibrated after their transformation into Discrete Composite probabilities to fit the twenty-nine play baseline, the logic of the exercise requires further serious explanation.

Having said this, it remains that it is ambiguous as whether the quotation refers to the false negatives at the macro level as I have considered them, or whether the quotation refers only at the micro level to the false negatives that occur among the forty-eight tests of Discrete Composite Analysis. This ambiguity should be clarified.

To return to the matter of the choice of boundaries for the discrete univariate analysis, the authors state that 'we find that no core Shakespeare baseline play has more than two individual rejections

(shaded aqua) in forty-eight tests, while no 'Claimant' play has fewer than ten rejections' (Elliott and Valenza, 2004, p. 338). While 'unwaveringly using machine-defined boundaries'[5] within the limits of two standard deviations from the mean for the twenty-nine play baseline, I found that *Much Ado About Nothing* and *The Merry Wives of Windsor* had five rejections each instead of one and two. They are plays with the highest proportion of prose with 87% and 72%, respectively (Wells and Taylor, 1987). Three plays, *Richard II, Hamlet*, and *Coriolanus*, had four rejections. Four plays, *Two Gentlemen of Verona, Othello, Cymbeline*, and *The Tempest* had three rejections. Although the increased number of rejections for the baseline plays does not alter their relationship with the fifty-one Claimant plays with ten or more rejections, it does alter their relationship with Dubitanda plays such as *Henry VI, Part 2*, with two rejections, *Henry V* with five rejections, *Henry VIII* (Shakespeare's part) with four rejections, and *Two Noble Kinsmen* (Shakespeare's part) with four rejections.

It is worth remarking that the correlation between rejection counts based on two standard deviations from the mean, 'unwaveringly using machine-defined boundaries', and the Continuous Composite probabilities for the twenty-nine play baseline is greater than that between the 'handfitted' rejection counts of Professor Elliott and the same Continuous probabilities of Professor Valenza.[6]

A fourth difficulty lies in the absence of an expressed consideration of the essential part played in the selection of a core baseline by prior probabilities from scholarly consensus. *The Tempest*, for example, with a Continuous Composite probability of 0.003689 is included in the baseline because of the unassailable confidence on the part of scholars and actors that it is uniquely the work of Shakespeare. In Bayesian statistical analysis, this prior belief would be accorded a probability estimated to be high enough to overcome *any* empirically derived evidence to the contrary. Only the *ad hoc* adoption of the Continuous Composite boundary set by *The Tempest* permitted it to remain within the framework of the 'clean, communized baseline'. Similarly, *Hamlet*, its four rejections with two standard deviations from the mean giving it a Discrete Composite probability of 0.013, must be included in a Shakespeare core baseline because of its defining quality as essentially Shakespeare. *Pericles 3–5* with a Discrete Composite probability of 0.06309 finds itself excluded from the core baseline, although its literary status as Shakespeare's would give it a Bayesian prior probability in excess of the arbitrary baseline boundary of 0.2316. Some discussion of the varied role of 'subjective' prior probabilities is appropriate.

References

Elliott, W. E. Y. and Valenza, R. J. (1996). And Then There Were None: Winnowing the Shakespeare Claimants. *Computers and the Humanities*, **30**: 211.

Elliott, W. E. Y. and Valenza, R. J. (2004). Oxford by the Numbers: What Are the Odds that the Earl of Oxford Could Have Written Shakespeare's Poems and Plays. *The Tennessee Law Review*, **72**: 323–453.

Fisher, R.A. (1947). *The Design of Experiments*. Edinburgh, London: Oliver and Boyd, 16.

Howson, C. and Urlach, P. (1989). *Scientific Reasoning: the Bayesian Approach*. London: Open Court, 8.

Jackson, MacD. P. (2006). The date and authorship of Hand D's contribution to *Sir Thomas More*. *Shakespeare Survey*, **59**: 69–78.

Merriam, T. (2005). *The Identity of Shakespeare in Henry VIII*. Tokyo: Renaissance Institute.

Taylor, G. (2008). Thomas Middleton. *Dictionary of National Biography*, Oxford: Oxford University Press.

Wells, S.W., Taylor, G., Jowett, J., and Montgomery, W. (1987). *William Shakespeare: A Textual Companion*. Oxford: Clarendon Press.

Notes

1 Words used by Professor Ward Elliott in an email to the author, 27 June 2008, with permission to publish. See also 'Moreover, unlike Discrete, Continuous is in no way *dependent on human judgment for making boundaries*' (italics mine), (Elliott and Valenza, 2004, p. 350).

2 This was a recommendation I offered several years ago.

3 Elliott and Valenza, 2004, p. 355. If the nine of the seventeen tests not included in the Clinic's consideration of time dependency are submitted to PCA, their first principal component correlates significantly with date of composition, -0.696 for $n = 29$.

4 Elliott and Valenza (2004, p. 338): 'When we count every test run, on every core baseline Shakespeare play, we find 98% true positives and only about 2% false-negative Shakespeare rejections. Applying the same process to fifty-one plays mostly by claimants produces 65% false-positive results. We ignore this result because false–positive results do less to prove a Shakespeare ascription than the false-negative results disprove it, and 35% true negatives, with a bit of aggregation, are more than enough to rule out fifty-one plays as Shakespeare's solo work'. Note that in this passage Dubitanda plays are not mentioned.
5 Words used by Professor Ward Elliott in an email to the author, 27 June 2008, with permission to publish.
6 Their correlation coefficients are respectively −0.76 and −0.64.

Appendix

Table A1

Play	Chronological order	Date given by Elliott	Fig. 1 Discrete	Fig. 2 Continuous	Fig. 3 PCA
R3	1	1593	1	0.898	0.200
Shr	2	1594	0.602	0.443	0.496
TGV	3	1594	0.602	0.062	0.417
Err	4	1594	0.602	0.423	0.224
R2	5	1595	0.232	0.034	0.124
LLL	6	1595	0.232	0.429	0.144
Jn	7	1596	1	0.135	0.078
MND	8	1596	0.232	0.674	0.135
Rom	9	1596	1	0.994	0.230
1H4	10	1597	0.602	0.952	0.197
Wiv	11	1597	0.232	0.013	0.575
MoV	12	1597	0.602	0.904	0.100
2H4	13	1598	0.602	0.613	0.213
JC	14	1599	0.602	0.464	0.276
Ado	15	1599	0.602	0.010	0.246
AYL	16	1599	0.232	0.398	0.214
Ham	17	1601	0.232	0.051	0.649
TN	18	1602	0.602	0.685	0.579
Tro	19	1602	1	0.847	0.670
MfM	20	1603	1	0.965	0.582
AWW	21	1603	1	0.884	0.709
Oth	22	1604	0.602	0.294	0.583
Lr	23	1605	1	0.907	0.901
Mac	24	1606	0.602	0.540	0.845
Ant	25	1607	1	0.959	0.896
Cor	26	1608	0.602	0.422	0.954
Cym	27	1610	0.602	0.995	0.931
Tmp	28	1611	0.232	0.004	0.973
WT	29	1611	1	0.973	0.930

Table A2

Play		Date	Fig. 4 Discrete	Fig. 5 Continuous	Fig. 6 PC1	Fig. 7 PC2
R3	S	1593	1	0.89791	0.839	0.800
Shr	S	1594	0.6018	0.44275	0.378	0.627
TGV	S	1594	0.6018	0.062283	0.533	0.673
Err	S	1594	0.6018	0.42266	0.693	0.813
R2	S	1595	0.2316	0.033822	0.891	0.786
LLL	S	1595	0.2316	0.42905	0.520	0.743
Jn	S	1596	1	0.13478	0.929	0.905
MND	S	1596	0.2316	0.67391	0.712	0.825
Rom	S	1596	1	0.99445	0.603	0.823
1H4	S	1597	0.6018	0.95223	0.692	0.901
Wiv	S	1597	0.2316	0.012563	0.385	0.621
MoV	S	1597	0.6018	0.90359	0.720	0.903
2H4	S	1598	0.6018	0.6127	0.817	0.792
JC	S	1599	0.6018	0.46432	0.563	0.779
Ado	S	1599	0.6018	0.009988	0.486	0.866
AYL	S	1599	0.2316	0.39824	0.797	0.888
Ham	S	1601	0.2316	0.050606	0.833	0.169
TN	S	1602	0.6018	0.68504	0.638	0.553
Tro	S	1602	1	0.84691	0.734	0.346
MfM	S	1603	1	0.96541	0.649	0.386
AWW	S	1603	1	0.8835	0.567	0.269
Oth	S	1604	0.6018	0.29374	0.671	0.420
Lr	S	1605	1	0.90747	0.705	0.114
Mac	S	1606	0.6018	0.53986	0.767	0.195
Ant	S	1607	1	0.95866	0.611	0.084
Cor	S	1608	0.6018	0.42183	0.825	0.040
Cym	S	1610	0.6018	0.99501	0.682	0.028
WPRZ	F	1611	1.09E–13	1.00E–15	0.063	0.592
Tmp	S	1611	0.2316	0.0036895	0.467	0.073
WT	S	1611	1	0.97333	0.642	0.044
VALN	F	1612	1.38E–12	1.00E–15	0.104	0.811
H8(Fl)	F	1613	8.44E–15	1.00E–15	0.046	0.242
H8(Jt)	F	1613	1.00E–15	1.00E–15	0.119	0.508
H8(Sh)	S	1613	2.77E–07	5.23E–07	0.798	0.013
MTOM	F	1613	4.36E–12	1.00E–15	0.060	0.253
CHNC	F	1617	1.00E–15	1.00E–15	0.015	0.597
LOYL	F	1618	1.00E–15	1.00E–15	0.014	0.645
DEMT	F	1619	1.08E–13	1.00E–15	0.032	0.269
BARN	F	1619	2.38E–12	1.00E–15	0.341	0.680
ISLN	F	1621	1.00E–15	1.00E–15	0.067	0.739

Table A3

Play	Author	Date	Fig. 8 Discrete	Fig. 9 Continuous	Fig. 10 PC 1
R3	S	1593	1	0.898	0.845
Shr	S	1594	0.602	0.443	0.518
TGV	S	1594	0.602	0.062	0.682
Err	S	1594	0.602	0.423	0.791
R2	S	1595	0.232	0.034	0.881
LLL	S	1595	0.232	0.429	0.821
Jn	S	1596	1	0.135	0.900
MND	S	1596	0.232	0.674	0.808
Rom	S	1596	1	0.994	0.698
1H4	S	1597	0.602	0.952	0.784
Wiv	S	1597	0.232	0.013	0.521
MoV	S	1597	0.602	0.904	0.829
2H4	S	1598	0.602	0.613	0.771
JC	S	1599	0.602	0.464	0.773
Ado	S	1599	0.602	0.010	0.610
AYL	S	1599	0.232	0.398	0.791
Ham	S	1601	0.232	0.051	0.420
TN	S	1602	0.602	0.685	0.646
Tro	S	1602	1	0.847	0.635
MfM	S	1603	1	0.965	0.693
AWW	S	1603	1	0.884	0.539
Oth	S	1604	0.602	0.294	0.718
PHOE	Mi	1604	4.335E–12	1E–15	0.063
Lr	S	1605	1	0.907	0.614
Mac	S	1606	0.602	0.540	0.670
MICL	Mi	1606	1E–15	1E–15	0.024
Ant	S	1607	1	0.959	0.521
Cor	S	1608	0.602	0.422	0.636
Tim	S	1608	4.355E–12	1E–15	0.369
Cym	S	1610	0.602	0.995	0.504
CHST	Mi	1611	1E–15	1E–15	0.024
Tmp	S	1611	0.232	0.004	0.446
WT	S	1611	1	0.973	0.572
NWIT	Mi	1613	1E–15	1E–15	0.024
MDIS	Mi	1615	1E–15	1E–15	0.026
WITC	Mi	1616	1E–15	1E–15	0.041
HENG	Mi	1618	4.355E–12	1E–15	0.246
WBWM	Mi	1621	4.355E–12	1E–15	0.033
GAME	Mi	1624	1E–15	1E–15	0.123

Table A4

Play	Author	Date	Fig. 11 Discrete	Fig. 12 Continuous	Fig. 13 PC 1
R3	S	1593	1	0.898	0.206
Shr	S	1594	0.602	0.443	0.509
TGV	S	1594	0.602	0.062	0.418
Err	S	1594	0.602	0.423	0.229
R2	S	1595	0.232	0.034	0.132
LLL	S	1595	0.232	0.429	0.130
Jn	S	1596	1	0.135	0.083
MND	S	1596	0.232	0.674	0.129
Rom	S	1596	1	0.994	0.231
1H4	S	1597	0.602	0.952	0.198
Wiv	S	1597	0.232	0.013	0.535
MoV	S	1597	0.602	0.904	0.090
2H4	S	1598	0.602	0.613	0.186
JC	S	1599	0.602	0.464	0.265
Ado	S	1599	0.602	0.010	0.212
AYL	S	1599	0.232	0.398	0.202
Ham	S	1601	0.232	0.051	0.608
TN	S	1602	0.602	0.685	0.541
Tro	S	1602	1	0.847	0.663
MfM	S	1603	1	0.965	0.544
AWW	S	1603	1	0.884	0.672
Oth	S	1604	0.602	0.294	0.541
Lr	S	1605	1	0.907	0.906
Mac	S	1606	0.602	0.540	0.845
Ant	S	1607	1	0.959	0.901
Cor	S	1608	0.602	0.422	0.955
Pericles 3-5	Q	1608	0.063	0.007	0.813
Cym	S	1610	0.602	0.995	0.929
Tmp	S	1611	0.232	0.004	0.972
WT	S	1611	1	0.973	0.927