Catching Reader Responses on the Fly

by Rosanne G. Potter[1]
Department of English
Iowa State University
Ames, Iowa 50011

Literary criticism, the art of making discriminating judgments or evaluations of literary works, depends in large measure on perspective. Reader Response Criticism approaches literary works from the perspective not of the writer, or the period, or the style, but from the perspective of the reader. One looks at interpretations of literary works and questions the knowledge base, interests, and psychological defenses of the readers who devised the interpretation. For more than 10 years, I have been using computational and statistical means to study linguistic and stylistic features in texts in an attempt to find and quantify textual controls in dramatic literature.

I conducted my first pragmatic study of real readers in the early 1980s but, at that time, did not concern myself with gender. In 1990 when I agreed to chair the Women's Studies Program at Iowa State, I began noticing differences between male and female student readers in my classes; I wondered whether males and females reacted to different cues or responded to the same cues differently and, most important, how I could catch the responses as the readers were reading. This paper describes my investigation of gendered reading; it begins with a brief summary of other pragmatic studies of readers - noting the presence, or absence, of sex and gender as factors in these studies.

Reader Response Criticism has been the least developed kind of criticism because, until quite recently, readers were thought to make only one contribution to the critical process: they were the sources of error, private associations, and misreadings. These negative judgments of readers flow from the assumption that there really is a "right" or "most complete" reading, and that, if properly trained, all readers can achieve it. This "right reading" assumption has undermined practical critics from I. A. Richards to Elizabeth Flynn and is still very much in practice in most American classrooms.

Although theorists like Wolfgang Iser have written book after book about "the reader," the researchers who have attempted pragmatic studies of real readers have been few and their methods painfully unsystematic. The first, and most famous, empirical study was performed at Cambridge University by I. A. Richards in the twenties. Richards' results were so devastating to him, and to many other, that for fifty years after, no one attempted to study the reading skills being learned in literature classrooms. In his 1929 book Practical Criticism, Richards reported the results of asking Cambridge undergraduates to read thirteen poems (authorship not identified and ranging from John Donne to minor poets), then to "comment freely" on their "readings." Richards documents — in excruciating detail — the many ways of misconstruing meaning in poetry when it is presented "without any hint of provenance"(5). Richards presents his findings not to indict the "products of the most expensive kind of education" (292), but to demonstrate that in all types of educational settings "we must cease to regard misinterpretation as an unlucky accident. We must treat it as the normal and probable event." (315) He rightly ascribed their poor readings to "bewilderment" (296) caused by the lack of "clues [about] authorship, period, school, the sanction of an anthology, or the hint of a context." (296)[2]

In her 1978 book The Reader, the Text, and the Poem: The Transactional Theory of the Literary Work,[3] Louise Rosenblatt reported on over twenty years of collecting student responses to unidentified poems; as I see it, she, like Richards, created abnormal test situations. In everyday reading, readers know the name of the author and can easily discover her/his dates, nationality, and school, and may have had their responses shaped by earlier readings of other works by the same writer or by earlier teachings about the writer or the work. By forcing readers to respond to the words of the text only, both researchers deny a reality condition in trying to create an unbiased test situation. This experimental design inevitably sets up perfect conditions for "errors" and encourages the discovery of differences between readers' responses.

Fifty years after the publication of Practical Criticism, Norman N. Holland in 5 Readers Reading decided to conduct "more or less undirected interviews with a few readers who had taken standard personality tests" (x).[4] He taped extensive interviews with five undergraduate students on Faulkner's "A Rose for Emily" and then read their responses to the story in light of their "identity themes" (56-62). Instead of finding a great deal of overlap between the readers, Holland found that the readers perceived very different stories. Holland arrived

at "four principles that describe the inner dynamics of the reading experience: "Style Seeks Itself" (113), "Defenses Must be Matched" (115), "Fantasy Projects Fantasy" (117), and "Character Transforms Character" (121). These principles are psychological descriptions of how readers transform the characters and events in stories to defend their own identity themes while reading fictions.

Holland, like Richards, used a free-response method, but Holland's method was molded by three interventions Richards had not allowed. The readers knew the author and the name of the story (some had even read it before in other classes). The readers' responses and personalities were both elicited by questions from Holland and reported by him. This method of gathering reader responses produces, in Holland's words, no "uniform core" of meaning "from the text" as opposed to "individual variations" contributed "from the people" (366). Thus, while Richards is distressed at the general decrease in ability "to make plain sense of poetry" (12), Holland explains "the way readers respond to literary characters as if they were real people" (xii) by applying Freudian terms ("transformations," "defenses," "fantasies") to the reading experience.

Although they perceive the outcomes of their experiments very differently, neither Richards nor Holland has a model that can be replicated, because neither has an experimental design with clear-cut categories for grouping responses. They inevitably emphasize differences among readers because each reader is treated as a separate case rather than has having features that can be clustered with other similar reader responses. In any study where readers respond "freely" and no attempt is made to identify features in the responses which correlate with features in the texts read, the results will inevitably emphasize difference.

My 1982 study of reader responses to the characters created in the first acts of 21 modern English-language plays<footnote text> could focus on agreement among readers because it asked all readers to respond to the same questions about seven character traits (dominance, intellect, excitability, speculation, poetry, education, attitude), using the same scale (e.g., markedly dominant, moderately dominant, neither dominant nor dominated, moderately dominated, markedly dominated). The reader responses were correlated to features in the language assigned those characters (like high or low use of questions, imperatives, fragments, exclamations, and seven other syntactic and/or semantic features). Since the research design asked specific questions and correlated the results with countable features in texts, there was no difficulty either in finding areas of statistically robust agreement among readers on character traits or in regressing the character trait data against the linguistic features to show which features figured at what levels in readers' judgments of character traits.

Literary scholars who do not know about dependent and independent variables or about objective methods of handling data, and who never attempt to gather qualitative information in quantifiable ways, are destined to discover, as Holland did, responses that have "nothing in common" (366).

Up through Holland, no particular interest in sex or gender as variables shows up. If women participate in the studies, that fact is either not noted or not considered a significant enough factor to require any balancing of the groups being tested. In the 1980s, the idea of gender-balanced samples begins to emerge, but since the general research methods continue to be highly subjective, the introduction of this variable hardly matters; the presence of sex as possible variant does not change the basic methods of analysis.

In Elizabeth Flynn and Patrocinio Schweickart's important 1986 book Gender and Reading, David Bleich reports on conducting an admittedly unscientific study of four females and four males (one of whom was himself) in an attempt to discover differences in male and female ways of reading male and female writers. Bleich's general conclusions are that males and females respond in similar ways to lyric poetry by male and female writers, but very differently to fictions. Men conduct a dialogue with the author about the characters and situations, while women enter the fictional world and allow themselves to see feelings more quickly than men do. This graduate-student pilot study established the theses to be tested in a larger, apparently scientific, contrast of the retellings of Faulkner's short story "Barn Burning" by 00 males and 00 females. Bleich, whose most famous book is entitled Subjective Criticism, does not use objective methods for handling the responses he collected. Instead he reads the responses subjectively and finds, not surprisingly, that they not only confirm his earlier findings, but also allow him to go on to even larger generalizations. The narrator's voice is the "mother tongue" and since separation from the mother is less significant for women than for men, women perceive men as "less-other" than men perceive women.

These assertions may be true, but we should not be misled into thinking that the use of objective methods to collect data means that Bleich's generalizations have any more truth value than if he had arrived at them without consulting any readers. Actually the data he collected and the conclusions are structurally unrelated. Starting from a gender-balanced sample does not necessarily prove anything about gender. Skill in research design is not widely distributed, especially among literary critics who have no training in even recognizing a well designed project when they read it.

In her essay on "Gender and Reading" in the book of the same title, Elizabeth Flynn describes collecting large numbers of responses from males and females (even though that was quite inconvenient at a mostly male school) and using random selection to achieve a balanced number of responses. Unfortunately, she also applies no objective tests to the data so carefully assembled. Like the male critics who preceded her, Flynn reads the essays and judges their adequacy as readings against her own critical standards supplemented by psychological terms describing human interactions. Examples of "domination" by or "submission" to texts are quoted and dismissed in favor of a "third possibility" in which "the reader learns from the experience without losing critical distance; reader and text interact with a degree of mutuality" (270). Flynn concludes by asserting that many males react to disturbing stories by "rejecting" them in an attempt to "dominate" the text, and females "more often arrive at meaningful interpretations of stories because they more frequently break free of submissive entanglement in texts" (285). Like Holland's, Flynn's conclusions are interesting and both sets of insights may have, as Oscar Wilde quipped, "the minor merit of being true"[5] but neither researcher has begun to prove anything. They have simply used a new method of establishing "authority" or ethical appeal.

At this point, I wish I could say and here, tah dah, is the reader response study that does what none of the others have done, but I come more to discuss the theoretical issues that are at stake when gender appears in empirical studies of reader's responses to literature, than to make final report on research.

Since my work on male and female responses to modern British literature texts is still under development and has only been described in print in a Belgian journal, I will describe it here in some detail. The study grew out of two occurrences in a 1990 Modern British literature course in which the students wrote reader responses to each assignment before class discussion. The first reader response, to Chapters 1 through 3 of Oscar Wilde's The Picture of Dorian Gray, surprised me because the readers' responses seemed to be sex-linked. Both male and female students commented on the "flowery" language and the ornate tone of the writing, but the males then asked, and I quote, "Is this guy queer?" or asserted "This guy must be a homosexual," while the females said "He is very sensitive" or "very poetic." In class, we discussed what they were reacting to and why some males drew conclusions that no females did. I gave the students the "sexual facts" on Wilde: that he was a married man and, as far as his biographers can discern, had not had any homosexual experiences at the time of the writing of this novel; that he did subsequently have such experiences and, five years later, was imprisoned for "gross acts of indecency," the 1890s' code term for homosexuality.

A month or so later, the second incident occurred. The students had read their first selection from Bernard Shaw's An Intelligent Woman's Guide to Socialism, Capitalism, etc. (Chapters 7-12). Here the difference was much more pronounced and intense: the female students related very positively to the text: they felt that Shaw understood the realities of women's lives and was arguing for improved economic conditions for women. A strong majority of the male students felt that Shaw was condescending to women and treating them as little better than children.[6] This difference erupted into a full-scale classroom confrontation and to the discovery that the males and females were in some cases using the same passages to prove their different positions. The strongest male and female speakers both wrote papers in support of their readings.

The female, who was simultaneously attending a senior seminar "Language and Gender," designed a questionnaire using selected passages (ones that "proved" her point, ones that "proved" the male student's, and ones that both asserted "proved" their opposing positions) and administered it to male and female student friends. Although the sample was small and neither stratified nor random, the results provided more anecdotal support for the proposition that male and female readers drew opposing conclusions from their readings of the selected passages. As a result of these two cases of striking gender differences in reader responses to literary texts, I designed an interactive reading experiment for use during the next offering of the same course.

The project I designed in 1990 and ran in 1991 goes back conceptually to Michael Riffaterre's 1959 insight ("Criteria for Style Analysis" Word) about the existence of places in literary texts that are commented on by almost all average readers (or ARs). Riffaterre asserts that disagreements (among critics about what passages mean) proves that stylistic devices (or SDs) have surprised readers. The unexpected use of language, according to Riffaterre, elicits interpretation.[7] I wanted to catch ARs responses to SDs in their first reading by inducing them to read new texts on a computer screen and to respond to anything that they found "surprising or unexpected."[8] My assumption about the best way to get a response (without interrupting the reading process) was to ask readers to take a simple action while reading, e.g., to "double-click" on words that seemed surprising or unexpected.

There were six readings: balanced by gender of writer (three females, three males) and balanced by genre within each gender (two works of fiction, one of non-fiction). The males (Oscar Wilde, Bernard Shaw, and James Joyce) were all from Dublin, though from differ-

ent classes and social spheres; the females (Katherine Mansfield, Enid Bagnold, and Virginia Woolf), were from London and Wellington, New Zealand, were all upper-middle class.[9]

The students self-selected into either MAC-lab readers and control-group readers. The MAC-lab readers came to the lab, entered a few demographic facts (their sex, age, major, and home state) in a logon procedure, and read for the first time an initial segment (chapter, or part of chapter) or a complete short work. As they read, they double-clicked on words, thus highlighting them and (although this was not spelled out to the students), simultaneously moving each highlighted word into a list tagged with the student's demographic facts.

After completing the reading, the students were asked to reply to six forced-choice post-reading questions designed to establish whether they (1) enjoyed the reading (i.e., did they want to read more by this author), (2) were experienced readers of texts like this one, and (3) felt competent (in terms of vocabulary and general comprehension) in the face of this text. They were then asked four expansion questions ("What is this work about?" What do you think of the writer?" "What do you notice as repeated?" etc.) and finally, they were asked to write a short paragraph of reaction to the reading.<footnote text> The control group read the same assignment before class, discussed it in small groups, and then wrote about it again afterwards. A contrast of the discursive writings by the MAC-lab and the control group students was anticipated but, unfortunately, not performed in the pilot stage.

The Mac-lab readers were overwhelmingly female (18 to 6). The ratio of females to males in this course is routinely 3 to ; the same ratio self-selected into the experimental group. The female/male imbalance was intensified when four males completed only five of the six readings, and one completed only four. I had hoped to find that males and females responded to many of the same words, and that some words were more surprising to females than males and vice versa. The results were so skewed as to be statistically unreportable; the best that could be said for the six lists of "surprising words selected" was that they showed as much variance within sex as between sexes.

This study of reader' responses, especially when it is seen in the context of other pragmatic and theoretical approaches to reading comprehension, can be refined in a number of ways. First, it needs to be conducted on larger samples of males and females which, as the research reported on in this paper shows, means moving out of the small upper-division classroom and into the large, always available, Freshman English pool. Second, more needs to be known about the readers; knowing their sex is not sufficient. If readers could be arranged along a scale of more or less "masculine" or "feminine," their responses could grouped to see if the social construct of gender is more useful than the biological differentiation into male and female. Third, the whole question of post-reading questions needs thorough re-thinking. Many students reported having very little memory of the texts they read on the MAC screens.[9] The students probably experienced some test and time anxiety because they knew they would have to answer questions after the reading; these anxieties may have interfered with their responses, their comprehension and subsequent memory of the texts.

Possibly, the most important re-design would be to pre-segment the texts, so that passages, rather than words, would identified by clicks. Word orientation tended to mean that "difficult" vocabulary items (including British spellings or usages) dominated the word lists. Passage orientation would group responses so that students who respond more slowly (at the end of a striking passage rather than near the beginning) would still be counted as responding to the same stimulus as the quicker, more experienced readers.[10]

This first of these improvements flows directly out of the comparison between the small and larger studies; it is prima facia clear that if one wants to study male versus female readers, the numbers of males and females need to be larger and more balanced. (Bleich's report on four subjects is used merely to investigate gender differences; the second study of 00 students is the one that is supposed to convince.) According to Mack Shelley, the statistician I worked with on the pilot stage of this project, I would need a sample size of at least 260 responses (balanced between males and females) to have enough degrees of freedom to start getting statistically significant results.

The second improvement has to do with triangulation, a research design achieved in my "Reader Responses and Character Syntax" project: the relation of two countable features through a third. In my 1982 essay, I counted occurrences of syntactic and semantic features and correlated them with characters who used these features, through readers' judgments of those characters' personality traits. Here, I counted the words and the sex of the reader, and probably should correlate these through the reader's scores on a standard test, like the Bem Sex Role Inventory. The Bem scores could be used to arrange the readers along a female/androgynous/male spectrum and might contribute to a better account of within-gender variance.

The third improvement, eliminating the post-reading questions, would keep the readers' attention focussed on the reading. The students were certainly distracted from selecting words as surprising, because they were concerned about whether they would have enough time to

complete the assignment in the class period. My study of gendered reading has just begun.

The Future for Studies of Gender in Reader Responses
When Reader Response Criticism is fully articulated, it will try to understand how readers' experiences, experiences tied to their class, age, ethnicity, and gender, affect their interpretations of literary works; it will try to assess the impact of information (or the lack of it), inclinations and disinclinations in readers. Unless testing techniques are defined that categorize kinds of responses (based on knowledge, personality traits, gender roles), unless reader demographics: sex, age, ethnicity, education are factored in, and until all factors are correlated to features in the texts, reports about the impact of gender on the reading of literature will be based on assertions, not upon research.

1 Presented at the IASSIST 92 Conference held in Madison, Wisconsin, U.S.A. May 26 - 29, 1992.

2 In her 1987 The Return of the Reader (London & New York: Methuen), Elizabeth Freund summarizes the "vices" as: "carelessness, self-indulgence and sentimentality [with] arrogance and obtuseness...not far behind." (32)

3 Carbondale and Edwardsville: Southern Illinois Press.

4 Holland had enunciated this model in his 1968 work The Dynamics of Literary Response (New York: Oxford University Press, 1968).

5 "Reader Responses and Character Syntax" in Computing and the Humanities, ed. Richard W. Bailey, (Amsterdam, New York, and Oxford: North-Holland).

6 The Writings of Oscar Wild ed. Isobel Murray, "The Critic as Artist, Part II," 278.

7 These responses formed 25% of the student's grade, were collected daily and returned at the beginning of the next class with brief positive reinforcement comments, i.e., no comments on writing problems, or "wrong" interpretations, just positive notes on insights, thought processes, and/or expression.

8 In her best selling 1990 book You Just Don't Understand, Deborah Tannen's descriptions of the differences in conversation styles may explain for the differing responses. Tannen asserts that women's conversational style emphasizes establishing community, while men's conversation style emphasizes competing for authority. The women students may feel recognized and valued when Shaw explains how the economic system takes advantage of the unpaid labor of wives and mothers. The men students may feel condescended to, put in the one-down position, when Shaw assumes the role of the authority explaining economic relationships to readers who do not understand the subject.

9 When many commentators mention a passage, regardless of whether they say the same things about it, Riffaterre asserts that they do so because the passage is surprising and calls for interpretation.

10 Relying on common language interpretations of these terms, I chose not to create a stipulative definition.

11 The texts were the first chapter of Enid Bagnold's Diary without Dates, "Bliss" by Katherine Mansfield, and the first chapter of Virginia Woolf's To the Lighthouse, Chapters 1 through 3 of Wilde's Dorian Gray, the first fifteen pages of James Joyce's The Portrait of an Artist as a Young Man, and Chapters 7 through 10 An Intelligent Woman,... Each file was approximately fifteen pages long in Microsoft Word.

12 This was the only required part of the post-reading responses (all others could be skipped); something had to be entered here to logout.

13 The texts were re-read when they came up in their normal places in the syllabus.

14 If I choose to pre-segment the texts, that would mean a complete re-design of the project. A taxonomy, like the one described by Teresa Snelgrove in her essay on George Eliot ("A Method for the Analysis of the Structure of Narrative Texts" in Literary and Linguistic Computing 5 [1990]: 221-225, of narrative-mode tags, possibly supplemented by persuasive-more tags could be employed to segment and pre-tag the texts; then the collected responses to any word within a segment could be accumulated to see whether narrative and/or persuasive modes and gender differences correlate. Differences in segmentation and/or tagging choices might also turn out to be gender-matched. If a number of male and female critics segmented and tagged the same texts, similarities/differences could be recognized and added into the variables checked in the responses of student readers. This augmented approach definitely merits consideration.