Using Sentiment Analysis to Ease Students toward or around Macroanalysis

Donald Secreast
Radford University

Today, digital humanities, particularly in the area of computational literary analysis, have become important enough that at some universities in America, graduate English majors who take a prescribed sequence of digital humanities courses can use them to satisfy a foreign language requirement. One of the most influential voices predicting new venues for digital humanities is Matthew L. Jockers. A list of his books provides a clear compass of Jockers' research interests: *Macroanalysis: Digital Methods and Literary History* (2013), *Text Analysis with R for Students of Literature* (2014), and *Anatomy of the Blockbuster Novel* (2016). Essentially, in *Macroanalysis*, Jockers argues that literary studies, in order to be relevant in the 21st century, must adopt a more scientific/computational approach: "Like it or not, today's literary-historical scholar can no longer risk being *just* a close reader; the sheer quantity of available data makes the traditional practice of close reading untenable as an exhaustive or definitive method of evidence gathering. Something important will inevitably be missed" (9).

Most helpful in understanding the mechanics introduced in *Macroanalysis* is a website that Jockers maintains, www.matthewjockers.net, in which he provides an ongoing dialogue of wide-ranging literary analyses using the statistical method explained in *Text Analysis with R for Students of Literature*. This website allows other researchers to share the results of their work

using Jockers' digital approach. The program "R" (derived from a social science algorithm) enables literary researchers to "mine" hundreds or thousands of texts using an algorithm that searches for specific words, phrases, or precisely defined structures. With compelling effectiveness, "R" enables researchers to gather information more quickly and thoroughly to create graphs based on such statistical products as word indexing, frequency lists, text-analysis comparison, stylistic analysis, and textual dating. Summarizing the underlying benefit of computational literary analysis, Jockers points out, "The approaches outlined in this book also have the potential to present contradictory evidence, evidence that challenges our traditional, impressionistic, or anecdotal theories" (*Text Analysis* viii).

While Jockers' method can generate almost limitless approaches to academic analysis of texts, he has also proven that his statistical technique has a more marketable application. For a short period, in association with Jodie Archer, Jockers formed a consultancy agency to "help writers, agents and publishers find the next books that everyone will love" (www.archerjockers.com). In their agency's website, they claimed to "analyze over 3000 data points" using their "advanced proprietary methods." These methods allow the agency "to predict bestsellers 80% of the time" (*The Writer Files*). It's interesting to speculate why this service was discontinued after a relatively short time. One can imagine Jockers and Archer receiving hundreds of manuscripts each month and having to deal with disappointed authors who have been told by an algorithm that their novels lack any possibility of becoming a blockbuster. It's bad enough to have a writer's workshop group tell you your manuscript lacks mass appeal, but how does a disappointed author respond to the disapproval of a computer who prints out three

thousand reasons why your novel is a flop and provides a detailed graph emphasizing the faults in your writing?

Why Privilege a Visual Approach to Literary Discussion?

From a teacher's point of view, the most attractive aspects of macroanalysis are the detailed graphs generated by the "R" program. Any analytical approach which provides visual aids in explaining and prompting discussions about the often abstract and elusive elements that constitute literary reality should be valued as a principal tool in the classroom. Certainly, other writers and theorists, from Gustav Freytag to Kurt Vonnegut, have developed their own charts and graphs to illustrate the different strategies and structures that make for compelling narratives; however, Matthew Jockers' method represents a genuinely new and purely digital approach to visualization.

No less enthusiastic than Jockers for exploring the possibilities of diagrammatic analysis in the humanities is Johanna Drucker who, in her book *Graphesis: Visual Forms of Knowledge Production*, provides an extensive history of graphs and a captivating argument for those of us in the humanities to recognize the influence our digital age is having on the way we think, the way we connect to the world and to other people—even the way we perceive. She asserts, "We need to develop a domain of expertise focused on visual epistemology, knowledge production in graphical form in fields that have rarely relied on visual communication" (iv). For those of us who have struggled with the task of introducing our students to critical thinking, Drucker's approach might appear at first to be just another sort of scaffold for analysis, another theory we

might try to stretch over our students' minds to filter out all the bad habits they acquired while being trained to pass their SOL's and other standardized tests in high school, learning to summarize rather than actually analyze, being trained to rally generalities rather than recognize significant details. She argues that graphs and graphic design represent a historic as well as an organic impulse for understanding that has deep roots in our human perception and cognition. Methodically, she provides evidence that, quite often, the visualization of emotion and abstract thought leads to an objectification, a concrete shareability, that enables us to more easily and accurately verbalize, organize, evaluate, and communicate those subjective states. In other words, for many of us in the humanities, visual structures clarify linguistic structures. As Drucker puts it, a graph "is a provocation to cognitive experience, but it is also an enunciative apparatus" (147).

Teachers who might be interested in exploring the use of visualization as a way to clarify and focus their own discussions of literary concerns or to encourage students to participate in poster presentations at conferences can acquire a thorough background in the history, philosophy, and relevance of graphic presentation in Drucker's *Graphesis,* particularly when she distinguishes between "visualizations that are *representations* of information already known and those that are *knowledge generators* capable of creating new information through their use. Knowledge generators have a dynamic, open-ended relation to what they can provoke" (65). Graphs infuse classroom literary discussions with the potential to expand in unexpected epistemological directions because by having visual constructs in front of them, students can more readily grasp the distinctions being made by other students.

Why We Don't Want to Start with Macroanalysis in the Typical English Class

Despite all the impressive results Jockers' research has produced, most English majors will likely be reluctant to embrace macroanalysis for three reasons. First, the traditional English major might be intimidated by the mathematics involved in Jockers' statistical analysis. Unless a course is specifically set up to provide extensive training in algorithms, a literature class simply won't have the time to acclimate English majors to the highly technical "R" program. My own personal experience with *Text Analysis with R for Students of Literature* was frustrating and took me back to those anxious nights when I was floundering in a programmed learning workbook for my trigonometry and calculus class in high school. And I ran into the same problem when I attempted to approach R through *R for Dummies*. Second, macroanalysis employs what might strike many English teachers, along with their students, as a counterintuitive approach to reading—what Jockers, under the influence of Franco Moretti, has come to call "distant reading" (*Macroanalysis* 17). While giving a historical, sociological, and geographical explanation of why traditional reading strategies are no longer adequate for contemporary literary theorists, Moretti provides a terse description of what constitutes distant reading: "Distant Reading: where distance, let me repeat it, *is a condition of knowledge*: it allows you to focus on units that are much smaller or much larger than the text: devices, themes, tropes—or genres and systems" (48-49). In contrast to close reading, macroanalysis, by way of distant reading, focuses on letting the computer do the reading for the researcher, running an extraordinary quantity of texts through

the R program then letting the researcher interpret the numerical results provided by its

algorithm.

Third, and perhaps most disturbing to a traditional English major, in permitting the

computer to have primary contact with the text, macroanalysis further distances the reader from

intimate contact with the text by intervening, to a certain extent, in the reader-response dynamic.

Fundamentally, statistical analysis is a form of structuralist criticism, and in that regard, it runs

the risk of leading to reductive consequences. In other words, even with all the possible

variations in the graphs generated by R, the statistical nature of the program leads to a sense of

attenuation, with the reader left outside the algorithm. When the algorithm becomes the primary

reader, then a significant amount of the response no longer belongs to the person; consequently,

an important connection to the text disappears, resulting in something along the lines of a

readout-reader relationship.

Modified Sentiment Analysis as Detour to or around Macroanalysis

No doubt, macroanalysis is already finding its way into the college curriculum and is

bound to establish itself as a traditional discipline, much the way literary theory established itself

as a parallel discipline to literary studies in the sixties and early seventies. To prepare students

for computational literary analysis, and to give them time to develop their math skills, teachers

can ease them toward or around Jockers' digital world by first introducing them to a less

alienating approach to analysis and visualization through a combination of "sentiment analysis"

and reader-response theory. Somewhat counter to Jockers' insistence that being "*just* a close

reader" puts English majors in a risky position, sentiment analysis, even with its goal of producing information-generating graphs, still requires students to engage in close reading of a given text.

In fact, with its fundamental reliance on a reader's reaction to specific elements in the text, sentiment analysis actually provides a framework for objectifying then visualizing those reactions. One of the most approachable and definitive discussions of sentiment analysis can be found in Bing Liu's *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. While sentiment analysis definitely has its own technical vocabulary (Liu 1-134), what students basically need to know is that it involves attaching numerical value to emotional responses— what Jockers will variously refer to as "scaled valences," "scaled sentiment," "sentiment," or "emotional valence" in the graphs found on his website. The more traditional term, established by earlier psychological researchers, is emotional valence, and it's the term I use when introducing my students to sentiment analysis.

Although Liu suggests that sentiment analysis in its contemporary form dates back to "early 2000," the importance of the concept goes back much further (xiii). Anyone who is a fan of *Mad Men* might remember the attention the ad men devoted to how potential customers might react to the way a product was presented to them—not to the product itself, but to its presentation. Far from being an abstract study of pure statistics, despite its use of quite sophisticated algorithms, sentiment analysis, as Liu points out, has become a pervasive presence in society: "Many large corporations have built their own in-house capabilities. Sentiment analysis systems have found their applications in almost every business and social domain"

(xiii). The recent Netflix documentary series *Social Dilemma* reveals just how extensively social media platforms rely upon applying statistical methods to our buying, viewing, and communicating habits. Liu describes this practice, formulated in sentiment analysis research, as working "with the aim of bridging the unstructured and structured worlds and facilitating qualitative and quantitative analysis of opinions" (xiii).

Trying to teach pure, statistically-informed sentiment analysis to English majors could be as counter-productive as dropping them prematurely into the deep computational waters of Jockers' macroanalysis. As a teacher, I chose to simplify my presentation of sentiment analysis by diluting it with the more cordial postulates of reader-response theory, particularly as developed by Wolfgang Iser. Basically, Iser insists ". . . the critic should not explain the text *as an object* but its *effect* on the reader" (*A Handbook of Criticql Approaches to Literature* 335). Even more pertinent to teachers who plan to introduce a seriously modified version of sentiment analysis to their classes is Iser's idea that ". . . a text does not tell readers everything; there are gaps or blanks, which he refers to as the 'indeterminacy' of the text" (355). At first glance, the connection between serious sentiment analysis and reader-response theory might seem extremely tenuous. However, if we focus on Liu's observation that sentiment analysis aims at "bridging the unstructured and structured worlds," and juxtapose it with Iser's assertion that a responsible reader's role is to bridge the "gaps or blanks" of a given text with his or her own responses, attitudes, or opinions in order to confront the "indeterminacy" of a text or part of a text, then what emerges is a methodology with which students can visually represent and more easily share their opinions as well as validate their subjective relationship with a text.
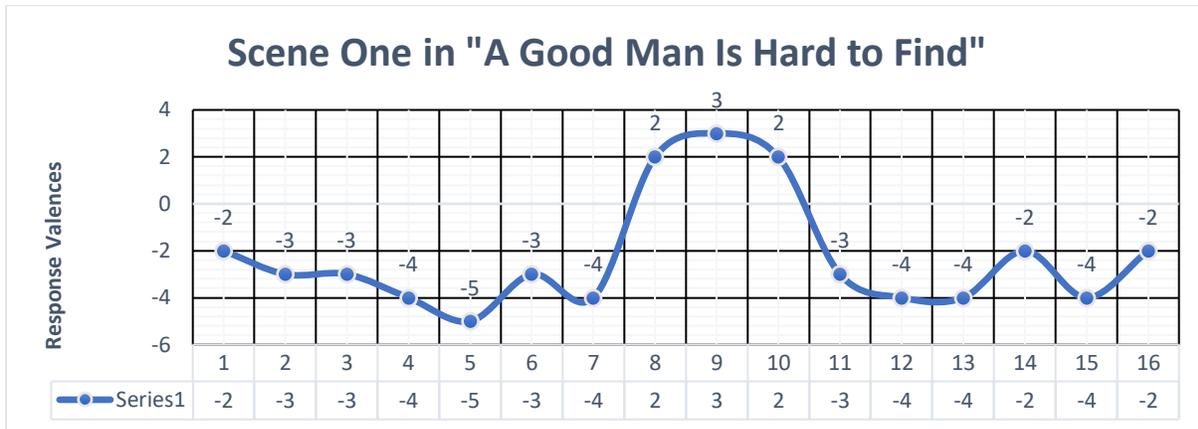
Consequently, teachers do not have to provide an extensive presentation on traditional sentiment analysis because all we're really taking from it is the idea of a group providing appraised accounts (which will become digitized) of their opinions or responses to an assigned reading, performance, incident, memory, or experience. In very strict applications of sentiment analysis, participants provide quite narrow responses, usually limited to what Liu calls the "5 sentiment ratings: *emotional negative* (-2), *rational negative* (-1), *neutral* (0), *rational positive* (+1), and *emotional positive* (+2)" (21). For Liu, these five responses, both emotional and rational, form the basis for sentiment analysis, particularly as used by product researchers. By limiting the range of numerical response from -2 to +2, traditional sentiment analysis aims to emphasize the comparative nature of its test audiences. In other words, sentiment analysis in its most statistical application isn't really concerned with individual emotional and rational responses as it is with establishing an easily quantifiable and general preference. Its intent is purely pragmatic and ultimately economic. Its numbers represent an attempt to clarify a product's likelihood of success.

In vivid contrast, sentiment analysis, when modified by a concentrated dose of reader-response theory, can take its participants in the opposite direction and clarify their subjective states for themselves and their classmates. Because forming a structured opinion or a thoughtful response to a piece of literature, music, experience, or any other encounter with the abstract requires a more complicated and personal evaluation than responding to a camera or protein shake, the assessment standards for this modified sentiment analysis must provide finer distinctions with which to map the students' interior landscape. Rather than just expressing a

positive or negative response, students of literature will have to be much more specific and develop agreed-upon number systems that reflect a range from, say, exuberance to misery or from euphoric to wretched. Other positive responses include cheerful, elated, gleeful, gratified, blissful, upbeat, engaged, buoyant, and satisfied. On the other hand, negative responses could include melancholy, anxiety, repulsion, conflicted, disturbed, frustrated, offended, depressed, bored, disappointed, and angry. In addition to choosing the spectrum of responses with which they will be working, students must also be tasked with prioritizing whatever lists they agree upon. This exercise provides them with the opportunity to broaden their vocabularies as well as sharpen their ability to make verbal distinctions.

One simple presentation I like to use when talking to students initially about using graphs to clarify rather abstract topics, particularly literary topics, involves a graph that I created to represent my response to the opening scene of Flannery O'Connor's short story "A Good Man Is Hard to Find" (Figure 1). I've allowed some redundancy in the numbering that appears on the graph just to make my references easier to follow. My sample graph is made up of four basic elements: 1) the y-axis, which provides the "Emotional Valences," or the intensity of the reader's reactions to the text; 2) the x-axis, which represents the specific points (technically called data points) in the scene to which the reader is responding; 3) the series 1 legend which provides the specific values for each data point and simply reiterates those values shown on the y-axis; 4) the line which connects the data points and provides a visual representation of the emotional wave evoked by the sixteen elements in the opening scene. Without much difficulty, the graph can be made more complex in a number of ways, most obviously by including more than one series, as

will be seen when we move to the discussion of a student graph produced by a group of high

school students in a Governor's School class back in the summer of 2016.

Scene One in "A Good Man Is Hard to Find"

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Series1	-2	-3	-3	-4	-5	-3	-4	2	3	2	-3	-4	-4	-2	-4	-2

Figure 1. Graph Representing Response Valences to Specific Points in the Opening Scene of Flannery O'Connor's Short Story "A Good Man is Hard to Find."

Scene one of O'Connor's story is slightly longer than one page. As is typical of an

O'Connor opening scene, it provides three important narrative necessities: 1) a definite hint of

the dramatic conflict, 2) an introduction of key characters, 3) and a strong suggestion of what our

attitude toward those characters should be. To a large extent, the opening scene provides a

thorough synopsis of the entire story. In collecting the information I planned to use as the data

points in my spread sheet, I began by reading the first scene, paying close attention to moments

or phrases or images in the text that evoked a definite response. As it turns out, O'Connor's first

scene prompted me to underline sixteen different phrases that called particular attention to

themselves. In the second step, I reread the underlined sections and tried to gauge what sort of

specific reaction was connected to each and how strong that reaction was. To each reaction, I assigned a number, either positive or negative. Some student readers might want to discuss the process of assigning numbers to their emotions, and this transition from reactions to numerical evaluations of those reactions must take into account what some students could see as the arbitrariness of such an exercise or, worse yet, as the process of dehumanizing one's emotions. Of course, even before the advent of macroanalysis or sentiment analysis, a certain type of English major always objected to analysis of any sort because it threatened to dissect the pure pleasure of reading. Once the teacher has provided the class with a suitable explanation of literary analysis's role as a tool for more deeply appreciating the art and craft that gives literature its power, attention can be focused on the mechanics of a modified sentiment analysis reading.

The implication of the conflict in "A Good Man Is Hard to Find" appears in the story's opening line: "The grandmother didn't want to go to Florida. She wanted to visit some of her connections in east Tennessee and she was seizing at every chance to change Bailey's mind" (117). My initial response of -2 to this conflict represents a pleasurable anxiety. I have a natural aversion to any sort of conflict, and that's why I find myself in a negative state of mind. However, because I know that a good story must have some degree of conflict, I appreciate the drama being presented so early. Consequently, while I'm negatively stimulated by the hint of conflict, I do acknowledge how smoothly O'Connor introduces it. This is the reason I rate my response as -2, an emotional state I would characterize as *anticipation*. Certainly, in a class discussion, some students will argue that they see anticipation as a purely positive response, but in such a situation, the teacher can introduce the idea of how emotions can be ambiguous. Even

anticipation of a positive event, like Christmas or a birthday, contains at its core a negative charge because to anticipate an event indicates that the event has not yet occurred.

Also contained in the story's first two sentences is my second response: a -3, mild disapproval, because a serious character flaw is also revealed: the grandmother's willingness to manipulate people in order to get her own way. Perhaps some students will argue that being manipulative is not necessarily a flaw, but I would point out that the grandmother's motive for manipulation is purely selfish. She wants her entire family to go to Tennessee—not Nashville or Memphis, but to east Tennessee. Students also need to be reminded that when this story was published, there was no Dollywood or Uber Gatlinburg. Perhaps most importantly, from the viewpoint of the grandchildren, east Tennessee has no ocean views.

The first view we get of the grandmother's son, Bailey, shows him completely ignoring his mother as he sits "on the edge of his chair" (117) reading the sports page. For different reasons, I respond to this introduction of Bailey with the same -3 (third response) I assigned his mother. His posture makes me think he is tense, perhaps because he is consciously resentful and avoiding the information his mother is trying to share. The fact that he chooses to ignore her rather than talk to her irks me a little more than witnessing his mother's manipulative tendencies. After all, most people, under the right circumstances, resort to manipulative behavior, but Bailey is being passive aggressive in this scene, and that sort of behavior is just another way to be manipulative.

The fourth moment in scene one which evokes an emotional response, a -4, occurs when the grandmother intensifies her manipulation of the family's vacation plans. Seeing that Bailey is

ignoring her, she presents proof that going to Tennessee is actually safer than going to Florida by pointing out to him the news story about the Misfit. This behavior increases my negative response because it reveals the grandmother's willingness to badger the family. If a -3 response corresponds to my feeling perturbed, then a -4 introduces a tint of anxiety. As soon as the grandmother goes into more detail about what the Misfit "did to these people" (117), my response (number five) drops to a -5 which represents a more fully developed attitude of anxiety. In contrast to his mother's intensified attempts at manipulation, Bailey responds to the threatening information with his default defense of simply ignoring his mother's scare tactics. Consequently, I find my sixth response rebounding slightly to -3, back to my attitude of mild, but engaged, disapproval.

Seeing that her scare tactics are having no effect on her son, his mother shifts her manipulative efforts to her daughter-in-law, further proof that she is definitely a shrewd and devious person who works hard to get her own way. This further development of her manipulative personality drops my reaction (response seven) to her to a -4, now slightly more negative than my attitude toward Bailey. In my eighth strong response—to O'Connor's description of the daughter-in-law—when I become aware of her comic intent, my attitude jumps to a +2 as a result of the phrase pointing out that the daughter-in-law's face was "as broad and innocent as a cabbage." The humor at this point provides well-timed relief from the mother-in-law's relentless manipulation and introduces a feeling of buoyancy. To anyone familiar with other O'Connor stories, this shift in narrative tone establishes an important element of her own manipulative technique. Even in her darkest stories, her sense of humor pervades. To readers

who hold the notion that humor can be inappropriate in tragic situations, O'Connor's comic

sense can be offensive. But anyone who understands her view of divine irony, that the most self-

righteous (who tend to be the most self-centered) deserve whatever punishment they eventually

incur, appreciates the justice of each inferno where such characters find themselves, infernos

designed by the characters' own attitudes and behaviors.

Not fully satisfied with portraying the daughter-in-law as a rather bland vegetable,

O'Connor further dehumanizes her by pointing out that the scarf she's wearing around her face is

tied "on the top like rabbit's ears"(117). Just why I find this comparison to an animal more

humorous than a comparison to a vegetable could be a topic for a class discussion, but I find my

response (number nine) rising to a +3 because my previous buoyancy is intensified to delight by

how O'Connor makes the logical but somewhat unexpected connection between a garden

vegetable and a rather opportunistic creature closely associated with the garden. Aside from the

rabbit being at the bottom of the animal food chain, perhaps just one step above the cabbage, the

comparison also implies that the young woman is more fertile than she needs to be. Interestingly,

this insinuation also serves a dual purpose by introducing the children to the reader. Being as

ignored by her daughter-in-law as she was by her son, the grandmother shifts her manipulative

intentions to the children.

Her insistence that going to east Tennessee rather than Florida will give her

grandchildren a chance to "see different parts of the world" reveals both her own narrow view of

the world but also the shabby quality of her machinations. In her own relentless way, O'Connor

establishes the grandmother as both pathetic and laughable. Just as the daughter-in-law emerges

as a less than human character, the grandmother also undergoes a process of dehumanization, a parody of the cranky, failed matriarch. Lacking the comedic surprise of the rabbit daughter-in-law, the shallowness of the grandmother still elicits a +2 response from me (number ten), partly because I can feel superior to her and her constricted view of the world. By eliciting this response of ironic gratification in her reader, O'Connor is already preparing her audience for being able to accept the massacre of the entire family by the Misfit and his two cohorts. Had O'Connor not insulated her readers against the violent ending of the story by making her characters unlikeable, the slaughter would be unbearable for most readers. For O'Connor, humor functions as a distancing device. Interestingly, the humor simultaneously allows the reader to be complicit in the cruelty.

Continuing to distance the reader from what should be the more innocent characters in the story, the children, O'Connor presents the son, eight-year-old John Wesley, as a brat—almost a miniature version of his grandmother. When he hears his grandmother complaining about going to Florida, he responds, "If you don't want to go to Florida, why dontcha stay at home?"(117) His disrespect prompts me to respond (number eleven) with a -3 upon hearing him this first time. However, my next response (number 12) of increasing disapproval drops to a -4 when John Wesley's younger sister June Star intensifies the tone of impudence by saying, "She wouldn't stay at home to be queen for a day."

Responding to her grandchildren's disdain, the grandmother resorts to the same scare tactics that she tried using on her son by summoning the threat of the Misfit. Her willingness to frighten her grandchildren, as unpleasant as they seem, resurrects my disapproval of her

manipulative personality. Consequently, my response (number thirteen) to her, at -4, matches my response to her grandchildren. But in a clear reflection of their family heritage, the two children demonstrate that they are completely contemptuous of the desires of other people. In response to his grandmother's asking him what he'd do if he met the Misfit, John Wesley insolently replies, "I'd smack his face"(118). This naïve belligerence comes across as both disturbing and depressing, especially if the reader is familiar with how the story ends. Where does an eight-year-old come up with the idea of smacking someone in the face? Possibly, it's an action he's witnessed or read about. And while his response implies a certain degree of violence, it also reveals that John Wesley really doesn't know what kind of deadly violence can occur outside his family. The ambiguity of John Wesley's place in the world O'Connor creates in this story makes me respond (number fourteen) with a -2, which reflects my anticipation of what's likely to happen to a character who isn't particularly attractive.

Similarly, when June Star resumes her rude assessment of her grandmother's need to go where the family goes, my disapproval of both grandchildren drops to -4 (response fifteen). "'She wouldn't stay at home for a million bucks,' June Star said. 'Afraid she'd miss something. She has to go everywhere we go'" (118). Voicing more resentment than her brother, June Star clearly ignores the grandmother's attempt at scare tactics, or perhaps her ignoring what her grandmother has just said about the Misfit indicates that she has already adopted her parents' total disregard for the older woman. Of course, in the final exchange of this opening scene, which prompts my response of -2 (number sixteen), a bemused irritation, when the grandmother refocuses her manipulative tactics by pointing out to June Star that she is the one who curls June

Star's hair, the granddaughter does respond to this comment—but perhaps only because it appeals directly to her own vanity. Because O'Connor merely summarizes June Star's response, "June Star said her hair was naturally curly," rather than providing direct dialogue, the close of this scene communicates the dismissive view the grandchildren have of their grandmother. But the humorous slippage between the grandmother's threat of the Misfit and curling June Star's hair keeps my reaction from being too negative.

By providing a method for close reading, the O'Connor graph, like many other sentiment/reader-response graphs, makes it easy for a teacher to demonstrate perhaps the most important lesson about the essential act of reading a serious text: to actually engage both intellectually and emotionally with literature lifts a reader out of a flat narrative flow and discloses the psychological undulations that enliven well-constructed art. Even students who are suspicious of analysis of any sort or who insist on being passive-aggressive in their reaction to literature can be encouraged to bring their particular biases into the graphing process, which is, after all, a basic tenant of reader-response theory. More advanced students might take the next step of distinguishing among which individual elements in the text generate specific responses. For example, a more complex graph can be constructed by tracking the responses connected to plot, character, action, setting, language, and theme. Obviously, such a graph would contain six series instead of just one. A more useful approach would be for the teacher to encourage advanced students to focus on just two or three narrative elements instead of graphing all six. However, even advanced students must be reminded that because a modified sentiment analysis approach to literary analysis can get overly complicated in direct ratio to the length and

complexity of the text under analysis, they should choose shorter rather than longer texts to analyze. Obviously, a short story is much more manageable than a novel. However, even short stories can generate fairly elaborate graphs. For example, "A Good Man Is Hard to Find" contains approximately thirty-one scenes. If each scene provides an average of say, sixteen responses which will be turned into data points, then a graph of the entire story would contain 496 data points—which would make an extremely long x-axis. Often, when discussing a complicated piece of literature or when class time for such a discussion is limited, I have found it helpful to assign students to groups who then analyze different sections of the text being graphed.

An Experiment with High School Students and Modified Sentiment Analysis

In the summer of 2016, I had the opportunity to focus on teaching a concentrated course in sentiment analysis to two groups of Virginia Governor's School students. I began by introducing them to the process of quantifying emotional states, and I was gratified to observe that these students were quite comfortable with the task of assigning numbers to how particular experiences made them feel. After all, on some level, most people do evaluate their emotional states by degrees of intensity. Ordinarily, the easiest scale to use when introducing students to rating emotions is -10 to +10 where -10 represents the lowest emotional state and +10 represents the highest emotional state. An interesting class discussion occurs when students are asked to numerically assess how a certain piece of music or poem or scene in a movie or chapter in a
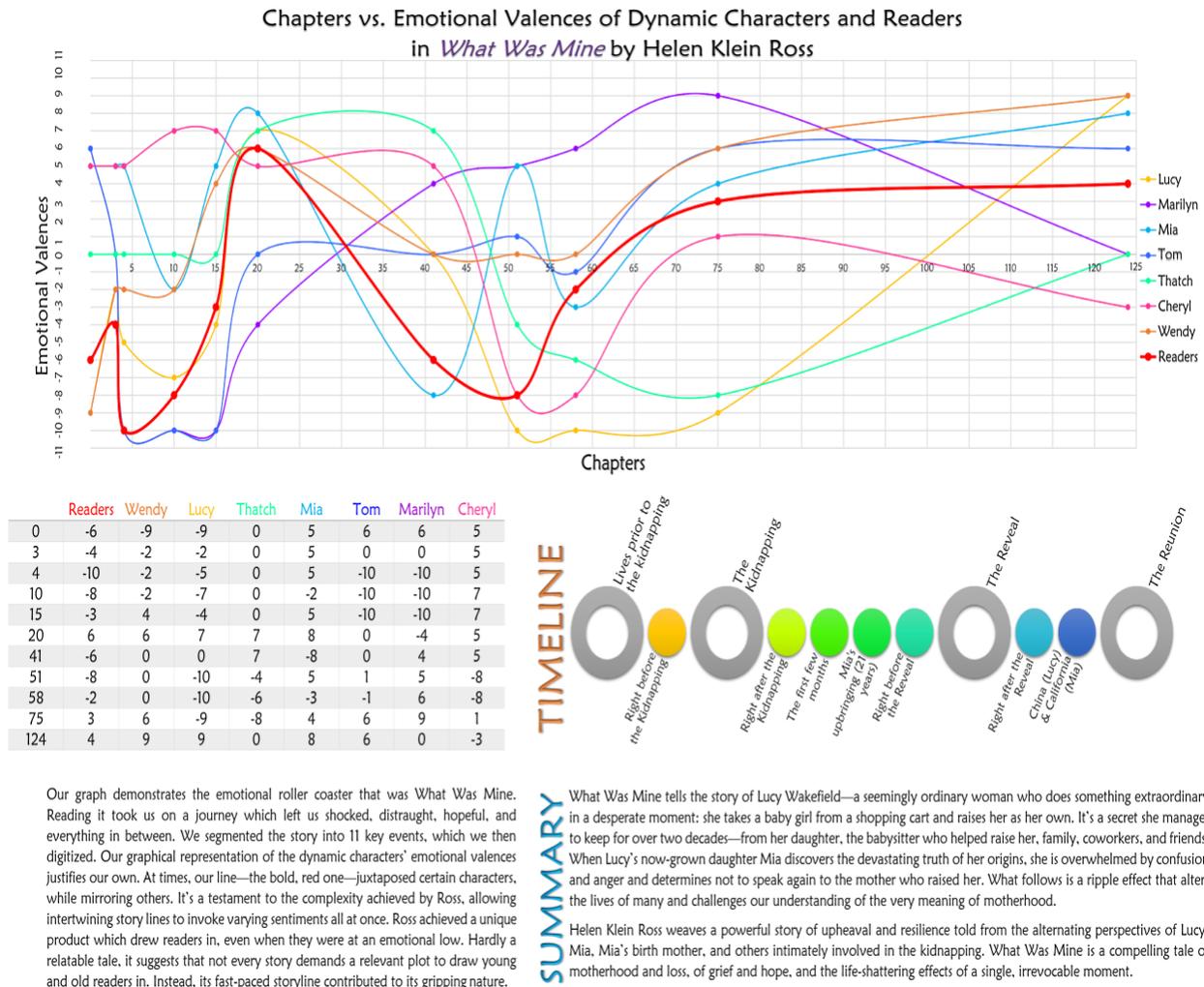
book has made them feel. Naturally, once they become accustomed to ascribing numbers to emotional states, they can apply the practice to a wide range of stimuli including speeches, social events, and academic situations. Once students have become accustomed to the practice of translating their abstract and emotional responses to literature into graphs, they will be less intimidated by the statistical transformations generated by Jockers' "R" program. Whether we call literary graphs knowledge generators or discussion generators, they aid in opening up literary analysis to what it should be: intellectual discovery through speculation, interrogation, and discussion.

After being reminded that the purpose of sentiment analysis is to create a visual representation of their mental or emotional journey, students must then be instructed in how to structure their digitized responses. In sentiment analysis terms, their responses will become data points, but these data points must be presented in a regulated sequence. As we've already seen, in analyzing O'Connor's short story, one of the most helpful sequential patterns is the scene. In a novel, the sequence might be chapters, or scenes within chapters. Once students have decided on what sort of pattern they want to employ, they are ready to isolate the elements in each scene that evoke definite emotions. These evocative moments become the data points to which the students will assign numerical values. Typically, the x-axis of the graph will show the chapters or scenes or timeline, and the y-axis will show the emotional valence. In other words, the x-axis typically tracks whatever structural concern the student has developed or discovered while the y-axis represents their emotional response to those structural or dramatic elements.

Once students have determined the structure of their analysis and collected their data points along that structure, they must then confront the most mathematical-looking aspect of this form of rudimentary sentiment analysis: creating a data base using their digitized emotions. The most convenient application for setting up such a data base is Microsoft Excel. Over the years, Excel has become more user friendly, and it can serve as a comfortable introduction to the more challenging programs like R. I discovered that my Governor's School students moved into Excel much more easily than I did and dressed up their graphs in ways I found completely captivating. As a matter of fact, I was so impressed by the students' graphic design skills that I neglected to stress the danger of their being much too ambitious in how broadly some of their projects started out. I would like to share a graph produced by one group of students after only three weeks of discussion and experimenting then provide a sample discussion of the graph to demonstrate how a visual representation of a student-produced sentiment analysis allows a class to more easily perceive the layers of a narrative, regardless of its genre. What English teachers should value the most when considering the advantages of using sentiment analysis to structure their discussions of literature is that the method can elicit more class participation and in the process sharpen students' analytical skills.

By graphing their responses to a piece of literature, students cannot avoid coming to terms with the "layering" that goes on in a piece of writing. Consequently, approaching class discussion through sentiment analysis can help students avoid getting stuck in a one-dimensional vision of the text under discussion. Because most high school and college English teachers usually require a novel as part of their literature component, I'm using a sentiment analysis graph

created by a group of students who had read Helen Klein Ross's *What Was Mine*. Although this

graph contains a few flaws which resulted primarily from my failure to adequately warn them

about the hazards of not limiting their analysis to a sufficiently narrowed text, I want to focus

primarily on how even a faulty graph can provoke constructive class discussion because it

provides an extremely effective visual presentation of how a novel can operate on multiple

emotional levels while following a chapter-by-chapter progression (Figure 2).

Chapters vs. Emotional Valences of Dynamic Characters and Readers in *What Was Mine* by Helen Klein Ross

	Readers	Wendy	Lucy	Thatch	Mia	Tom	Marilyn	Cheryl
0	-6	-9	-9	0	5	6	6	5
3	-4	-2	-2	0	5	0	0	5
4	-10	-2	-5	0	5	-10	-10	5
10	-8	-2	-7	0	-2	-10	-10	7
15	-3	4	-4	0	5	-10	-10	7
20	6	6	7	7	8	0	-4	5
41	-6	0	0	7	-8	0	4	5
51	-8	0	-10	-4	5	1	5	-8
58	-2	0	-10	-6	-3	-1	6	-8
75	3	6	-9	-8	4	6	9	1
124	4	9	9	0	8	6	0	-3

Our graph demonstrates the emotional roller coaster that was *What Was Mine*. Reading it took us on a journey which left us shocked, distraught, hopeful, and everything in between. We segmented the story into 11 key events, which we then digitized. Our graphical representation of the dynamic characters' emotional valences justifies our own. At times, our line—the bold, red one—juxtaposed certain characters, while mirroring others. It's a testament to the complexity achieved by Ross, allowing intertwining story lines to invoke varying sentiments all at once. Ross achieved a unique product which drew readers in, even when they were at an emotional low. Hardly a relatable tale, it suggests that not every story demands a relevant plot to draw young and old readers in. Instead, its fast-paced storyline contributed to its gripping nature.

SUMMARY

What Was Mine tells the story of Lucy Wakefield—a seemingly ordinary woman who does something extraordinary in a desperate moment: she takes a baby girl from a shopping cart and raises her as her own. It's a secret she manages to keep for over two decades—from her daughter, the babysitter who helped raise her, family, coworkers, and friends. When Lucy's now-grown daughter Mia discovers the devastating truth of her origins, she is overwhelmed by confusion and anger and determines not to speak again to the mother who raised her. What follows is a ripple effect that alters the lives of many and challenges our understanding of the very meaning of motherhood.

Helen Klein Ross weaves a powerful story of upheaval and resilience told from the alternating perspectives of Lucy, Mia, Mia's birth mother, and others intimately involved in the kidnapping. *What Was Mine* is a compelling tale of motherhood and loss, of grief and hope, and the life-shattering effects of a single, irrevocable moment.

Figure 2. Chapters Versus Emotional Valences of Dynamic Characters and Readers in *What Was Mine* by Helen Klein Ross.

Fundamentally, this chart of Ross's novel provides the necessary elements of a complete sentiment analysis: 1) the graph, which provides a visual image of how readers and selected ("dynamic") characters respond to important dramatic situations in the novel, 2) the data base, which provides the numerical values connected to the y-axis, the "emotional valences" being

measured and the chapters which establish the movement along the x-axis, 3) a legend or

"timeline" which provides a more specific explanation of which events in the designated

chapters give rise to the responses being measured on the y-axis, 4) a "conclusion" which gives a

brief discussion of how their data points highlight "intertwining story lines" in the novel, and 5)

a "summary" of the particular events in the novel to which the students are responding. While

working within the parameters of required components, this group still manages to introduce a

creative element into their analysis by graphing not only *their* emotional responses to the major

events of the novel but also by graphing what they interpret as the emotional responses of

selected "dynamic" characters in the novel to those same events. So, instead of having just one

series making up their data base, as seen in my O'Connor graph, this graph provides eight series.

In the summary section of the students' chart, basic plot information is provided to clarify

the events to which the sentiment analysis applies. The main character, recently divorced Lucy

Wakefield, obsessed and in a deep, sustained depression because she is unable to have a child of

her own, kidnaps an unattended infant in an IKEA and for twenty-one years raises the child,

whom she names Mia, as her own. Lucy gives her a privileged upbringing, hiring a Chinese

woman, Wendy, as her nanny for fifteen years. Parallel to Lucy's and Mia's lives, the novel

follows the suffering and transforming recovery of Mia's biological mother, Marilyn

Featherstone. Eventually, Lucy becomes a successful novelist, achieving fame after ghost-

writing a fictionalized version of her kidnapping. In a somewhat hard to swallow series of

coincidences, Marilyn uncovers the truth behind Lucy's novel and tracks down her kidnapped

daughter through Facebook. Understandably, Mia is seriously stunned by the revelation that

Lucy is not her mother but is, in fact, a criminal who stole her from another family. Traumatized by Mia's anger and facing possible arrest for her crime, Lucy flees to China where she will eventually move in with Mia's former nanny, Wendy, and her family. In the final chapters of the novel, Mia and Marilyn have agreed to forgive Lucy and try to help her avoid prosecution.

Timeline, Legend, and Spreadsheet

The "Timeline" created by the students, the data that constitutes the x-axis of the graph, provides two interrelated sets of information. On the right side of the chart appears a sequence of eleven circles which represent the major dramatic events, the structure of the story. Among the eleven events which the students will be responding to, clearly four of them, represented by the larger, gray rings, carry much more dramatic weight than the other seven: 1) the characters' lives before the kidnapping, 2) the kidnapping, 3) the "Reveal," the climactic section wherein Marilyn confronts Mia with the truth about Lucy, and 4) the reunion at the end of the novel when Mia has come to China to meet with Lucy.

On the left side of the "Timeline" an actual spreadsheet appears, detailing three sets of information: 1) the specific chapters to which the group and selected characters are responding, constituting the x-axis, 2) the numbers indicating the intensity of their responses, constituting the y-axis, and 3) the specific names of the responders designated by separate columns which identify the individual emotional valence points in the actual graph. To further aid in identifying which valence belongs to which responder, a color-coding system links the names on the

spreadsheet to the names in the legend on the right side of the graph. This is an automatic function of the Excel application.

The legend section of the graph lists and color-codes seven characters' responses as well as the group's or "official readers'" response to the eleven dramatic events that structure the novel. That's why their red line is noticeably thicker than the other seven colored lines. So in a class discussion with limited time, we might start by comparing how the other seven responses differ from the "official" response. In an actual classroom setting, a large part of the discussion would focus on the group presenting their graph, explaining what narrative strategies prompted each of their emotional responses and why they assigned the specific numeric values that they did to those responses.

What impresses me at this point with the group's analysis is that they are demonstrating some awareness of how necessary it is to provide specifics in a literary discussion. When I taught critical thinking to freshmen and mid-level English majors, I found myself spending much more time than I expected explaining to them the difference between analyzing and summarizing, and, more frustratingly, the difference between generalities and specifics. This graph, however, provides 1) specific characters who are responding in 2) specific degrees to 3) specific dramatic situations. In addition to demonstrating an ability to find a specific structuring method in the novel and narrowing the range of characters they will discuss, this group has also engaged in a hypothetical empathy that in itself adds a definite degree of intellectual engagement with the text. In the course of constructing a graph of their own emotional responses to the novel, the group has projected itself into the emotional possibilities of seven characters. This strategy

establishes a second layer of class-wide discussion because other students in the class can not only question the data points connected to the "official" readers but also question the official readers' representation of the characters' responses. While sharpening students' analytical skills, this experiment in hypothetical empathy also nurtures their ability to connect with a variety of personalities that they might never meet in their actual social spheres.

Through their group discussion, the students agreed on how the characters' responses should be gauged specifically in relationship to their own official reader response. Later on, we will see that the students come to recognize that their own responses are influenced by 1) how they extrapolate the selected characters' responses to the same dramatic situation, while at the same time, 2) carrying echoes of the other characters' emotional valences along with them from previous chapters, and 3) anticipating what the next plot point will do to certain characters. Such a discussion requires an ongoing comparison and adjustment sequence in which the students must keep distinguishing between how the dramatic situations made them feel and how those situations made the seven individual characters feel. So even on this speculative empathetic level, the group is engaged in arriving at psychological/literary distinctions. For example, if we follow the visual wave pattern formed by connecting the data points provided by the official readers plotted along the x and y axes on the graph, through the eleven plot points, the novel has provoked the emotional equivalence of the "roller coaster" ride described by the students in the conclusion section of their chart.

Viewpoint: How the Antagonist Becomes a Protagonist (Chapters 1, 2, 4, 10, and 14)

Clearly, they begin by sympathizing with the infant's kidnapper, Lucy, as the first two chapters establish her despair over not being able to have a baby of her own and before she commits her crime. She has become so obsessed with her desire to be a mother that her husband has left her. The "official" emotional valence is a -6 out of a possible -10. (These boundaries can be extremely flexible, depending on how widely the students think the emotional responses should range and what sort of vocabulary they agree to use to describe their spectrum of responses.) However, when Lucy is in IKEA, in chapter 3, and sees an infant, apparently abandoned, she feels protective, then possessive. Sensing Lucy's blossoming awareness of possibility, of motherhood, the official readers respond with a slight rise in emotion, from -6 to -4. However, when the readers are faced with the actual kidnapping, their response drops to the lowest valence provoked in the whole novel, a -10, what might be labeled as shocked disapproval.

Skimming from chapter 4 to chapter 10, the official readers register a -8 response to the self-justifying explanations that Lucy provides to herself and her readers. Giving rise to an engaging class discussion, this group should be ready to explain how Lucy, after kidnapping an infant, has slowly begun to win sympathy from her audience. Most likely, some class members will not agree that Lucy deserves the small rise in sympathy registered by the official readers, but a more general question might help explain how we can begin to identify with a bad character in literature or in movies. Similarly, this point in the graph might also give rise to a discussion

regarding the role that viewpoint plays in any piece of literature. As a matter of fact, because this graph allows for an ongoing comparison of eight different viewpoints, this group should emphasize that viewpoint is definitely a major device which allows for the layered structure of a well-written narrative. During this discussion, the observation should be made that even the "official" reading is actually made up of a multi-layered response; the members of this group first produced individual data points then averaged the points before plotting their graph. Similarly, the data points representing the seven characters' responses are hypothetical responses, so some members of the class could raise the question about how trustworthy such responses might be. Now would be an appropriate moment to discuss the importance of being an empathetic reader.

As the next five chapters (10-14) indicate, Lucy presents herself as a dedicated and fond mother. So persuasive is her voice, the official readers grudgingly grow even more sympathetic by responding across this stretch of narrative from -8 to -3. To justify this significant rise in their attitude toward Lucy, the group should present specific comments made by Lucy which mitigate the reprehensible circumstances through which she became a mother. Students should realize that behind Lucy's possibly self-serving comments, the author is also employing narrative strategies to make her appear more sympathetic. The rest of the class should express their interpretations of why the official readers are allowing themselves to be swayed. This part of the discussion can also introduce the vital digital concept of "granularity" in a graph and reemphasize the importance of specific references to the text.

The Relationship between Details and Granularity

Generally speaking, in discussions of a graph, granularity refers to the number of data points that constitute the analysis: the higher the number of data points, the higher the granularity. In "Valence Is a Basic Building Block of Emotional Life," Lisa Feldman Barret refers to "emotional granularity" when observing, "Some people made categorical distinctions, characterizing their experiences in discrete emotion terms, whereas others characterize their experiences in broad, global terms (37)." This different approach to describing one's emotional valence should be discussed by students to make them aware of their own method of evaluating how they respond to specific stimuli. On a graph, the degree of granularity is also dependent upon how many structural divisions appear along the x-axis. For example, in the graph of *What Was Mine*, the students have provided a total of eleven data points for seven characters and one official reader which totals eighty-eight data points. At first, this number sounds like a fairly high granularity. However, if we take into account that only eleven situations in a 124-chapter novel have been responded to, then the granularity of this analysis diminishes considerably. In terms of literary analysis, what this concern with granularity highlights is the importance of details, particularly in the form of direct quotations, images, and specific scene references.

The most apparent weakness of the chart comes after chapter 75. The students provide no data points until the final chapter—leaving forty-nine chapters without *any* responses. In sentiment analysis terms, the last half of the graph provides no "granularity" in its analysis, a weakness that could have been avoided if I had been less dazzled by their design skills and more

emphatic in my directions about narrowing their topic. Notice in the first five chapters, the students provide three data points for a total granularity of 24. This reading provides a comparatively finer granularity of analysis compared to the zero data points across those last forty-nine chapters. It's very clear that something happened to the students' attention to the text toward the end of their analysis. Maybe time ran out either for reading or for responding. This problem provides another important discussion point which should be covered either in a conference with the group or in class as a learning moment. Unlike actual macroanalysis which is specifically designed to deal with vast amounts of literary data, sentiment analysis works more effectively when applied to smaller literary fields. Despite various shortcomings of the graph, noticeably after chapter 75, the group's analysis still offers ongoing revelations about the author's skill at manipulating emotions, even in the artificial context of "averaged" responses and hypothetical responses.

Even a Faulty Graph Reveals Manipulation (Chapters 15, 16, 17, 19, and 20)

Returning to a consideration of the graph's strengths, I want to emphasize that the students have done a good job of isolating the main dramatic situations from chapter 1 through chapter 75. However, in discussing the granularity of their graph, the class might explore the effect that their need to summarize such broad sections of the dramatic structure has on the comprehensiveness of their responses as well as the responses they ascribe to the novel's "dynamic" characters. For example, by the time the readers reach chapter 20, when the stolen

child, Mia, has turned 21, the readers have finally crossed from their negative response to a very positive view of the kidnapper. At +6, the official readers have achieved what will be *their* highest state of mind for the entire novel. The six chapters that have been condensed in the two data points constituting the line of emotional valence between chapters 15 and 20 cover Lucy's efforts to fabricate a plausible narrative for how she became an "adoptive" mother and her ongoing paranoia about her crime being discovered.

Just how well Lucy succeeds in deceiving everyone is reflected in the two chapters, 17 and 19, in which her sister, Cheryl, admits to how happy she was for Lucy but also confused by some of her behavior, though having no idea of her sister's duplicity until her crime is revealed twenty-one years after the kidnapping. Perhaps what might be most disturbing to other readers, those who pay specific attention to Marilyn's narrative in chapter 16, is how much she suffers as she recalls those first desperate days following her infant's kidnapping. In a class discussion of this chart, some students might want to know why Marilyn's hypothetical response registers as only a -4. At the end of her chapter, she must face the disheartening fact that the police have given up on finding her daughter. Certainly, we can understand that after the shock she's experienced earlier in chapters 4 through 14 where she was assigned three consecutive -10 responses, she might be entering a less intense emotional state reflecting a stunned acceptance that she has lost her child. However, even if the official readers assume that Marilyn is now "numb" instead of shocked, they might resort to a thesaurus to come up with a more definite emotional state that corresponds to that -4. To assume that Marilyn is simply numb seems to imply that she isn't feeling any specific emotion at this point. To be numb would reflect more of

a 0 on the graph than a -4. Perhaps a phrase like "morbidly resigned," or "defeatedly acquiescent," or "complicitly fatalistic" would clarify what state of mind is indicated by a -4.

On one level, as the official readers moved through the narrative developed in chapter 20, they were paying closest attention to Lucy's voice. Because she establishes herself as a responsible and loving "mother" who is determined to provide her daughter with a privileged upbringing, our readers are ready to give her some approval. After all, as Lucy and even Marilyn remind us, Marilyn was neglectful as a mother. Apparently, the official readers are willing to take a narrator's opinion about herself or himself at face value, regardless of the implications vibrating beneath the admission. So, as Marilyn is feeling guilty and regretful about herself, Lucy rises above her understandable paranoia and begins to feel pretty good about herself, with a valence of +7. If we count from Marilyn's -4 up to Lucy's +7, we see that Lucy's state of mind is actually +11 higher than Marilyn's. What we might be seeing in this chapter is an underlying dynamic in reading. Even if a character has committed a terrible crime, but in all other areas he or she presents herself or himself as basically decent with good intentions, the reader can be persuaded to tolerate or even approve of that character.

Fostering Hypothetical Empathy through Close Reading (Chapter 41)

Adding to the positive aspects that Lucy herself presents as justification for her behavior, the official readers also ascribe hypothetically positive responses to six of the other seven characters throughout chapters 15-20. Only Marilyn through this section of data is assigned negative responses. Consequently, as a class, we might consider the possibility that even a

seriously criminal behavior can be forgiven if 1) the criminal has other good traits, 2) a majority of other characters, who are good people, approve of the criminal, 3) the victim of the crime shares some degree of responsibility for the crime, and 4) the criminal herself speaks in a persuasive enough voice. Reading then, the moderator of the discussion might point out, is not a simple linear movement through a plot. What's being clearly revealed so far by interpreting the graphic response to the novel is that characterization and voice can strongly influence how we respond to the overall narrative movement regardless of plot details.

The next data point after chapter 20 isn't recorded until chapter 41. Once again, in a class discussion or in a conference with this group, the question should be raised as to why they skip twenty-one chapters before responding to the narrative. As already pointed out, in a long text such as a novel, students will have a difficult time providing a satisfactory granularity for their graph because of the overwhelming density of the narrative details. Because the emotional valence in chapter 41 has dropped to -6—a significant decline from the +6 assigned to chapter 20—the group should be prepared to explain what specific events through the elided twenty-one chapters influenced their response. Understandably, chapter 41 could exert a negative impact on the reader because it is told from Marilyn's point of view. She has come across the novel *Baby Drive* which Lucy has helped write. Although Lucy's name isn't on the novel and some of the details of the kidnapping have been changed, Marilyn still has the premonition that through this novel, her stolen daughter is reaching out to her. So, on one level, a reader should be feeling a little more positive. The twenty-one-year-old mystery is about to be solved. Most likely, because the official reader has now developed such a strong empathy for the criminal explains how a

drop of twelve valance points has occurred.

Apparently, a great deal happens over the next twenty-one chapters because the next data point the official readers record at chapter 41 has plummeted to -6. Clearly, because twenty-one chapters are compressed into a single set of responses from five characters and the official readers, the granularity of this section of the graph is suspiciously low. Even the data points that are provided in chapter 41 appear confusing. For example, a data point is ascribed to a fairly minor character, Thatch, who is Mia's half-brother from the remarriage of Marilyn and whose viewpoint occurs only once, in chapter 82. Yet, in chapter 41, a data point with a +7 valence appears under his name. Equally confusing, a data point of -8 is assigned to Mia who, like her half-brother, is not heard from in any of the intervening twenty-one chapters that we assume are being summarized in chapter 41. For the most part, all of the *indirect* references made to Mia in these twenty-one chapters, either from Lucy's, Wendy's, or Cheryl's points of view, reflect a bright, privileged, perceptive child growing into a confident, accomplished young woman. Necessarily, in a class discussion of the emotional valence data points at chapter 41, the official readers will need to explain 1) why their response has dropped so noticeably and 2) why their interpretation of Mia's state of mind has become so negative.

Graphing as a Tool for Discovery (Chapters 40 and 41)

Perhaps most confusing about the data points appearing in chapter 41 is the fact that the official readers record Lucy's emotional valence at 0. First of all, Lucy is the main viewpoint in the novel. In the twenty-one chapters that build up to the data point at chapter 41, Lucy's voice is

heard in nine of the chapters. The other characters who appear in more than two chapters are Marilyn and Cheryl, both providing their viewpoints three times. To Cheryl, the official readers assign an emotional valence of +5, and to Marilyn, they assign an emotional valence of +4. Certainly, it's possible for a character to be neutral in her response to a dramatic situation, but Lucy is so emotionally invested in her own narrative that many students in discussing this section of the sentiment analysis will have serious questions about that 0 data point. For example, at the end of chapter 40, when Lucy is preparing to attend a book signing for *Baby Drive* in San Francisco, she expresses her concern that Mia's biological mother lives only thirty-three miles from that city. She tries to reassure herself that such a distance should provide a reliable buffer between her and running into Marilyn. Nevertheless, despite her rationalization, her last comment in chapter 40 is "Her proximity—or my sense of her proximity—made my insides liquify" (131). This is a powerful image for the anxiety Lucy feels, and it is strong enough to push her out of any emotion even remotely related to neutrality.

Possibly the official readers have simply let their power of empathy slip. Or they might be allowing some other set of narrative devices to influence their own response. If they are being influenced by their feelings about Lucy, then they need to explain how they can extrapolate their -6 response from her 0 response. On the other hand, if they're being influenced by the emotional valences they've registered for Cheryl and Marilyn, they must explain how the +4 and +5 responses drag the official reader's response all the way down to that -6. The only other negative data point, of -8, has been assigned to Mia. Clearly, such a low state of mind could pull the official readers down to that -6, but as we've already discussed, this group will have to explain

why Mia's emotional valence is so low when, for twenty-one chapters, most references to her imply only positive values.

One possible explanation the official readers might give is that they are anticipating the "Reveal" section of the novel which they establish as occurring in chapter 51. After all, in chapter 41, Marilyn has pieced together enough of her research on kidnapping and information provided in Lucy's novel to suspect, on a compelling intuitive level, that the story does establish a connection between her and her kidnapped daughter. In the final paragraph of chapter 41, Marilyn is on her way to the Barnes & Noble where she plans on meeting the author of *Child Drive*. Still, if the official readers are recording a -6 valence for themselves because they're worried about the damage that is about to be done to Mia's relationship with Lucy, they must also explain their -8 evaluation of her state of mind since Mia's viewpoint is completely missing from the twenty-one chapters condensed into the chapter 41 data point. Although the readers can definitely feel their valence drop as a result of their anxiety about what is *about* to happen, they must clarify, by exercising their empathetic response and consulting a thesaurus, that -8 for Mia. Strictly speaking, if they are projecting their anxiety onto Mia, for the sake of accuracy, they should restrict their deeply negative valence to their own set of data points and not confuse their anxiety with Mia's, who cannot be feeling any such anxiety because her viewpoint is limited to only what she can know and not to what the official readers can anticipate. In other words, to anticipate a character's future state of mind should not be confused with the reader's own response to that character's future state of mind.

Granularity and the Friction of Viewpoint (Chapters 41, 46, and 51)

Because only ten chapters are condensed into the data points found in chapter 51, as contrasted to the twenty-one chapters condensed between chapters 20 through 41, the granularity represented by chapter 51 is considerably higher than the granularity represented by chapter 41. The official reader's response at this data point has fallen to a -8, down from the -6 recorded in chapter 41. Now, if we adhere to an interpretive tendency that was discussed earlier, that the official readers' reactions are being influenced by which particular character's viewpoint they have been sharing, then the -8 valence experienced by the official readers can be justified. All of the Lucy chapters throughout this section of the novel depict the truth of her kidnapping becoming exposed, until finally, in chapter 46, she comes home to find her apartment "ransacked." Although she has experienced increasing paranoia after her encounter with Marilyn at the book reading, once she gets back home and a few weeks pass without the police appearing at her door, Lucy has almost convinced herself that she has escaped being exposed by Marilyn. Unknown to Lucy, however, Marilyn has made contact with Mia, becoming friends with her on Facebook. Eventually, following Marilyn's request, Mia takes a DNA test and finds out that the likelihood Marilyn is her biological mother is 99.9%. Earlier in the novel, when Mia was a child, Lucy had told her that she was adopted. To Mia's shock, though, the story Lucy told her does not match the account she hears from Marilyn. The ransacking that Lucy discovers is the result not of a burglary, as she first assumes, but of Mia searching for evidence of her true origin.

In class discussion, the group needs to explain why they give a -10 valence to Lucy in

chapter 51. First of all, the rather coarse granularity of their graph at this point conflates at least four major emotional moments experienced by Lucy. First, she feels excitement because she is appearing as a reader at her first book signing. Second, she feels increasing paranoia when she encounters Marilyn at that reading. Third, she feels some relief when, after several weeks, nothing seems to come from her brush with Marilyn. Fourth, she feels a shock when she discovers her ransacked apartment. During their discussion, the official readers should explain why these four reactions add up to a -10. They have to keep in mind that in chapter 58 when Lucy actually confronts the reality that Mia knows how she came to be Lucy's daughter, the emotional valence they assign Lucy at that point is also a -10. This equivalence prompts an interesting consideration for a full class discussion: the shock of thinking your apartment has been burglarized can be as powerful as the shock of knowing your daughter has found out you kidnapped her when she was an infant. The class might want to discuss the possibility of extending the positive and negative ranges of their y-axis. Or they might actually decide that the emotional value of suspecting you're the victim of thieves is equivalent to being exposed as a thief.

Granularity and Ambiguity (Chapters 47, 49, and 51)

Another compelling discussion question that should come up in class is why the official readers in chapter 51 assign both Mia and Marilyn the same valences of +5. In chapters 47, 49, and 51, what is presented through Mia's point of view is the deterioration of her life with Lucy. Following a steady downward movement, Mia's three chapters develop a state of mind that

begins with the statement, "Getting that message really made me freak" (146), which refers to the first contact Marilyn makes with her through Facebook. After going through an understandable period of denial, during which she questions Lucy about the details of her adoption, sensing a few faint irregularities in what Lucy tells her and what she remembers from the *first* time Lucy told her about her adoption, Mia makes the decision by the end of her conversation with Lucy to confirm the friend request that Marilyn had sent to her. Getting over her initial shock, Mia studies Marilyn's Facebook page and acknowledges to herself that she definitely does share facial features with at least two of Marilyn's other children. She also realizes that Marilyn could not have been the fifteen-year-old unwed mother from whom Lucy claimed she adopted Mia.

When Mia works up the courage to message Marilyn, she responds almost immediately, sending her links to newspaper articles about the kidnapping and also telling Mia that her real name is Natalie. Still not completely convinced by the end of chapter 49, Mia speculates that Marilyn could be crazy, and besides, she doesn't like the possibility that her life isn't at all what she thought. Now, these two chapters establish a deep emotional and psychological struggle within Mia. So far, what I expect the rest of the class to be reacting to would be the doubt that has entered Mia's attitude about her identity and what emotional valence might be assigned to Mia given the questions she now faces about her own origin. The class might question the fairly comfortable valence of +5 the official readers have assigned to her at the chapter 51 data point. Just from the standpoint of the spectrum of emotions implied by the -10 to +10 this group has chosen to work with, a +5 translates to a state of mind beyond "buoyant." I would encourage the

class to discuss the emotional dynamics that they have encountered when dealing with moments in which they have struggled with their own self-doubt.

Prompting further discussion regarding Mia's +5 emotional valence at chapter 51 is the dramatic action she takes in that same chapter. She had been contacted by Detective Brown, who was in charge of the case from the very beginning, when she was kidnapped from the IKEA. He is the one who first asks her to take the DNA test. Initially, Mia is reluctant to take the test, but then, after a fight with Lucy, she decides to cooperate with his and Marilyn's request. Responding to the test results, Mia declares, "February 1, 2012. The day that changed my life forever"(157). Possibly, the official readers might argue that Mia's statement reflects a positive response to the revelation about her true origin. Such an argument should provoke an interesting discussion about Helen Klein Ross's use of ambiguity. Quite possibly, all of her character's speak ambiguously. This aspect of the novel's dialogue could provide its own insightful discussion. If Mia's declaration can be taken as a positive response to what she's discovered, we should still ask the official readers to explain why that +5 reflects her emotional valence. But as we've already discussed, this group might explain their data point as representing an average of Mia's emotional states through the ten chapters condensed in chapter 51.

However, just in chapter 51, Mia experiences a variety of dramatic moments that call into question such a positive interpretation. Most significantly, toward the end of the chapter, when Mia searches through Lucy's closet for some confirmation that the DNA test might be wrong, she discovers an old suitcase which contains undeniable evidence that Lucy has kidnapped her. Finally persuaded that Marilyn has been telling her the truth about Lucy, Mia describes her state

of mind: "The closet walls seemed to close in and I felt as if I were about to be crushed"(161). Convinced that she has to escape the home that now makes her feel "as if bugs were crawling all over me" (161), she is so distraught that she has trouble locking the door behind her, and it seems to be saying to her "*You don't live here anymore*" (162). These three highly charged dramatic moments indicate that Mia is experiencing 1)an extreme sensation of claustrophobia, so strong, in fact, that she feels physically threatened, 2)formication, or parethesias, a form of "tactile hallucination," which is that sensation of insects crawling on one's skin, often caused by intense anxiety (*healthline*), and 3)a sense of alienation so pronounced that she has lost some degree of muscular control. Certainly, this traumatized state, this last view we have of Mia in this ten-chapter conflation, casts doubt upon the suitability of Mia's +5 emotional valence at this point.

Slippery Empathy and Ambiguity (Chapters 51, 53, and 58)

When they reach chapter 58, the group's emotional valence has slid up to -2, although Lucy is still suffering at a -10. Their response might be on the rise because they are now identifying more closely with Mia's birth mother, Marilyn, while retaining a diminishing sympathy for Lucy. Mia's emotional valence has now dropped from the +5 in chapter 51 to a -3—a full eight points. In the seven intervening chapters between chapter 51 and chapter 58, Mia's viewpoint occurs only once, in chapter 53 which picks up immediately after Mia's departure from her kidnapping mother's apartment. She is in tears and her sense of alienation grows even sharper because in leaving behind the apartment building where she grew up, she becomes more deeply aware that she is leaving behind the life she has known for twenty-one

years. Definitely, the official readers are correct to assign a negative value to Mia's state of mind. But considering the questions that were raised about why they evaluated her valence at +5 at chapter 51, the class might ask them to explain why she has experienced such an emotional drop when the alienation she is feeling in chapter 58 appears to be very similar to her alienation in chapter 51. As she rides in a cab across New York City, Mia continues to reevaluate her relationship with Lucy, whose deception makes Mia doubt the worth of her entire existence. That -3 doesn't seem to represent the psychological weight of such an identity crisis. Students should share their ideas on how such a crisis would make them feel. In addition, Mia cannot think of where she can go or with whom she might share her sense of loss. Her boyfriend is out of the country, and she isn't prepared to reveal the details of her condition to her best friend. More than likely, this sense of isolation compounded with her sense of alienation would have to pull her emotional valence lower than a -3.

In their defense, the official readers can point to the end of chapter 58 when Mia contacts her favorite high school teacher, Christine Laniere. Not only will Ms. Laniere provide a sympathetic ear, but she has always let her students know that her home can be used as a haven. At this moment in chapter 58, the official readers can reasonably argue that Mia's mood could be higher than we might expect, given that she finally believes she has found an avenue of relief from how she has been betrayed by the woman who has been the center of her world for twenty-one years. Nevertheless, the final image in chapter 58 shows Mia ignoring a phone call from Lucy and saying, "I didn't have a mom anymore" (171). Without a doubt, Mia has found some relief in knowing that she can share her pain with Ms. Laniere. At the same time, this relief could

counterbalance her loss of stability and identity enough to rate her emotional valence at -3.

As can be gathered from all the data points connected to chapter 58, Marilyn is obviously the only character who has an emotional valence in the positive range: +6. While such a number implies that now the official readers, with their -2 valence, are being influenced primarily by all of the other characters *except* Marilyn, the class might want to consider just how strongly those other negative valences are affecting the official readers. After all, their -2 valence is considerably more positive than Lucy's -10 and Cheryl's -8. A serious question could be raised about that valence attached to Cheryl. In the seven chapters condensed into the chapter 58 data base, Cheryl's viewpoint doesn't even appear. The graph doesn't clarify from where that -8 response for Cheryl actually originates. The last time Cheryl's viewpoint appeared was in chapter 35. Here, we seem to have another instance in which the official readers engage not only in hypothetical empathy whereby they try to apprehend a character's emotional valence but rather they are confusing their own valence with what should be an approximation of that *character's* emotional state. In such instances, the readers' conjecture is turned toward themselves rather than toward the character. Consequently, what should be an exercise in extrapolation becomes, instead, interpolation. As students become more familiar with the intellectual discipline of close reading, they should become more practiced at recognizing when their interpretation slips into interpolation. Naturally, an important discussion can take place if the official readers are offered the chance to explain how they may have extrapolated Cheryl's drop in emotional valence from +5 to -8 over those twenty-three chapters between 35 and 58 without having the benefit of reading Cheryl's viewpoint.

Granularity and Anticipation (Chapters 56 and 58)

Considering, then, the fact that the official readers' valence is comparatively higher than Lucy's and Cheryl's, perhaps the group can make a case that their -2 does reflect a positive influence coming from Marilyn's +6. In both chapter 56 and chapter 58, where her viewpoint again appears, her emotional valences are understandably on the rise. She makes first voice contact with Mia in chapter 56. Then in chapter 58, she breaks the news to her family that she is about to go to New York to meet her kidnapped daughter. As we have seen in several previous instances, many responses to a novel do develop as a result of anticipation, either when the readers anticipate implied future events in the plot or when they anticipate the effects those events will have on certain characters. Therefore, in addition to practicing their skills of empathy, serious readers also find themselves developing a capacity for inventive anticipation, both of which represent creative acts on the part of the reader. Specifically in chapter 58, Marilyn finds herself in a pure state of expectation, a condition which gives the official readers a chance to evaluate the dual nature of literary anticipation: 1) how it allows an author to dramatize a character, and 2) how that dramatization moves the reader. With this consideration in mind, the class should discuss how +6 accurately reflects Marilyn's level of expectation.

To properly initiate a discussion of this question, the teacher needs to once again emphasize to the class the importance of providing sufficient granularity in sentiment analysis. As was mentioned in the beginning of this essay, the underlying power of Matthew Jockers' macroanalysis approach is his algorithm's ability to sort through literally hundreds of texts, often

collecting the number of times a single word or phrase appears. A closer look at chapter 58

reveals that Marilyn's anticipation is more complex than simply being excited about meeting her

kidnapped daughter. Definitely, that specific excitement is a strong current pushing Marilyn's

emotional valence, but several other currents are also shaping her mood. Not surprisingly, her

first response after making contact with Mia is to contact her first husband, and Mia's biological

father, Tom. To Marilyn's surprise, Tom's reaction to the news of their daughter being found is

much more subdued than Marilyn thinks it should be. He gives her a vague excuse for not being

able to come immediately and reuniting with their daughter. His suggestion that mother and

daughter need to conduct their initial reunion without his presence in order to better "bond"

doesn't ring true to Marilyn, and regarding Tom's fatherly capacity, Marilyn concludes, "But

Tom is a limited person. He is just doing the best he can" (185). Considering this

disappointment with her ex-husband's reaction, Marilyn's emotional valence might not be as

high as +6. Students should be encouraged to discuss their own experiences with how they have

reacted when other people close to them have not shared their excitement over some event. The

tendency of most people in this situation is to momentarily question the validity of their own

excitement. Or their frustration with close friends or family who didn't share the same degree of

excitement must have dampened their own enthusiasm at least for a short period of time.

　　　After Marilyn has contacted Tom, she must inform her new family about being reunited

with her lost daughter. Of course, her present husband, Grant, and their three children know

about Marilyn's kidnapped daughter, but despite her joy over the prospect of being reunited with

Mia, Marilyn does harbor a slight misgiving about how her children will respond to the news that

a new member is about to be added to the family. To be more precise, Marilyn must deal with how *each* of her three children will respond. Her basis for doubt is a fear that one or all of her children might feel alienated or displaced somehow by Mia's joining their family. Underlying her concern for how her children will respond to Mia is the fear that Mia might not accept her new brothers and sister. Marilyn also carries a deeper doubt that Mia could even refuse to visit her new family. After all, Mia has been raised in rather luxurious surroundings in New York City while Marilyn and her family live in a much less sophisticated home. In light of these fears, the class discussion might focus on whether or not Marilyn's emotional valence would be as high as +6, especially with the understanding that she is still carrying her disappointment from her conversation with Tom. Many in the class might argue that despite all of Marilyn's doubts, her underlying excitement about reuniting with Mia can certainly push her emotional valence up to +6, maybe even higher. A discussion regarding the power of doubt to undermine the strength of a person's hope provides an opportunity for students to examine their own experiences when they have had to deal with such simultaneous, conflicting emotions.

Fluid Viewpoints: Irony and Anxiety (Chapters 59 and 75)

Following another lengthy abridgement, seventeen chapters, the official readers arrive at their tenth set of data points with chapter 75. Understandably, the plot has moved over a great deal of territory. Most importantly, Lucy has fled to China to avoid being prosecuted for her crime, and Mia has joined her new family in San Mateo. Through the seventeen compressed chapters, the narrative's primary focus has been on Mia, whose viewpoint occurs seven times;

Lucy, whose viewpoint occurs five times; and Marilyn, whose viewpoint occurs four times. In data point seventy-five, the official readers register their emotional valence as +3, up noticeably from their -2 in chapter 58, suggesting that their empathy has shifted from Lucy to Mia and Marilyn. Class discussion could begin by asking the official readers to explain why they have begun to identify more closely now with Mia and Marilyn. Or some members of the class might want to ask why the official readers' response isn't even higher than that +3. In other words, Marilyn's +9 emotional valence should be exerting a more positive influence on that +3. In contrast, others in the class might insist that Lucy's -9 should be exerting a more negative influence on the group's response. An especially notable feature that occurs in the graphs of reader-response/sentiment analysis is a marker for irony—when two diametrically opposed data points appear, as in this instance when Lucy's emotional valence is -9 and Marilyn's valence is +9. Looking back over all the data points, students will see that this ironic relationship between the two characters is a fairly regular property of the graph. Students might suggest the possibility that Marilyn's and Lucy's valences cancel out each other in the official readers' response.

One apparent reason the official readers can give for their response of +3 in chapter 75 is that they are now identifying most closely with Mia and her emotional valence of +4. Considering this connection between the reader and Mia, the class will need to discuss how accurately that +4 measures Mia's emotional valence. In the seven chapters which reveal Mia's viewpoint, she struggles with 1) excitement about reuniting with her biological mother, 2) her anger with Lucy, 3) her sense of loss, in terms of her personal history, 4) her anxiety about joining a strange family, 5) her inability to function when she returns to college, and 6) both the

uneasiness and pleasure in shifting her identity from a city girl to a country girl. In exploring Mia's turbulent ambivalence about being introduced to her biological family, students might observe how Helen Klein Ross most effectively presents one of her major themes: dramatic human emotions seldom arise from a single cause or operate as a pure motivation. As active readers, students can't help but discover the psychological layers that create a believable and compelling fictional character.

When Mia is about to meet Marilyn for the first time, her anxiety is apparent in the difficulty she has in trying to choose what clothes she will wear to the airport. In responding to this device of characterization by action, class members might share their own experiences about feeling that kind of nervousness. Even if they're preparing for an event which will be enjoyable, they might consider how that nervousness affects their overall emotional valence at that moment of preparation. Considering that the official readers have ascribed a +4 to Mia in chapter 75, they should clarify if she experienced that same valence in chapter 59. Another recurring theme emerging this far along in the novel highlights the role that anxiety and anticipation play in the characters' lives. The class might discuss whether or not these two states of mind play such a prevalent role in their own lives as they also explore what seems to be the relationship between these two psychological states.

Oddly, when Marilyn does appear and Mia can already recognize physical similarities between her biological mother and herself, her response is far from joyous when she confesses: "I felt I was falling down a hole with no bottom" (190). This sensation is reminiscent of the disorientation she felt back in chapter 59, when she discovered that Lucy *had* kidnapped her. In

addition to her nervousness at meeting her biological mother in chapter 59, Mia also has to deal with increasingly desperate texts from Lucy. A close discussion of this chapter might lead students to speculate that this is possibly the most critical moment in Mia's development as a character.

For the rest of the chapter, Mia numbly accepts her role as Marilyn's daughter, but clearly she isn't yet comfortable with it because as the chapter ends, she admits that she still feels "awkward" with her new identity, and that in reality she and Marilyn are as much "strangers" to each other as they are to the other people standing around the baggage carousel (191). Most telling is Mia's focus on the "dark hole" (191) of the carousel, which clearly reprises that image of the dark hole introduced in the chapter's opening. Just from this chapter, the class can make a strong case against that +4 valence presented by the group. However, if those official readers are arriving at their chapter 75 emotional valence by averaging all seven chapters where Mia's viewpoint appears, they can argue that she gradually overcomes her anxiety and not only adapts but comes to appreciate the transformation of her identity fostered by her new family and exposure to country living. Ultimately, the class should try to decide if what Mia gains from reuniting with her new family outweighs what she realizes she must give up.

Insufficient Granularity as Analytical Impediment (Chapters 72, 75, 79, 120, and 124)

After chapter 75, the graph suffers its most consequential impediment because no data points are provided by the official readers until chapter 124, which means that this final set of

data points represents a compression of forty-nine chapters. Strictly speaking, such a lengthy extrapolation implies more of an elision, more of an omission which strongly suggests that the line of the graph running from chapter 75 to chapter 124 is more likely to be tracking interpolation instead of extrapolation—generalization instead of analysis. Rather than spend a tremendous amount of discussion time speculating about what actual emotional valences were accumulating through each of the elided chapters between chapter 75 and chapter 124, class time will be better spent simply focusing on the final set of data points given for chapter 124. The difficulty of trying to discuss a sentiment analysis that lacks even a minimum of data points—not to mention *any* data points—gives the instructor the opportunity to once again emphasize the importance that granularity, specificity, plays in creating a persuasive graph.

Except for Lucy's sister, Cheryl, whose emotional valence is -3, Mia's half-brother, Thatch, whose emotional valence is 0, and Marilyn, whose emotional valence is also 0, all of the other characters, including the official readers, register positive valences. If students haven't already discussed the significance of a 0 emotional valence, now would be an excellent time to explore what state of mind corresponds to a 0. Some students might equate such a null reaction to the once popular attitude of "Whatever," or to the more political response of "It is what it is." In the three chapters featuring Cheryl's viewpoint, 72, 79, and 120, she first experiences the humiliation of being "a kidnapper's sister" (236). In her view, and it's a view she's certain the rest of society shares, the scandal has tainted her husband and children as well. When Cheryl finds out from Mia that she and Marilyn do not plan to press charges against Lucy and that Mia is actually going to China to bring Lucy back "home" (311), Cheryl's dominant emotion is

anger. She deeply resents the fact that Lucy has always succeeded in avoiding responsibility. With obvious bitterness, she reiterates the seriousness of her sister's crime against Marilyn and Tom: "She took their baby. She took their marriage. She took the lives they were meant to have" (312). Given the depth of Cheryl's anger and resentment and her refusal to forgive Lucy, despite Mia's and Marilyn's willingness to forgive, surely, some class members will disagree with the official readers' evaluation of -3. However, the official readers might maintain the accuracy of their response by pointing out that *exactly because* Mia and Marilyn have begun to forgive Lucy Cheryl's valence isn't more negative than that -3. Along a similar line of argument, the official readers might defend their own +4 valence at chapter 124 by arguing that their response isn't higher, even in light of Lucy's +9 and Mia's +8, *exactly because* to a certain extent they share Cheryl's disapproval of Lucy's crime and the fact that she is about to get away with all the damage she has done to a child, a marriage, and a family.

What I hope most clearly emerges from this discussion of the students' graph is how it compels them to become close readers who must engage with textual details in order to recognize the intricate dynamics of literary devices and narrative strategies, becoming more adept at comprehending and communicating the layered and fluid texture of any narrative. The graph becomes an interface, a visual field of transactional reader-response observations, a tangible meeting and departure point that not only connects them more firmly to the text and their responses but also to the responses of their classmates. In addition, students will feel more confident in formulating their insights arising from addressing specific textual details, structures, and stylistic strategies. By objectifying students' responses to a piece of literature, reader-

response/sentiment analysis graphs allow class discussion to flow more freely because questions and disagreements can be directed at the graph and less directly at the students whose interpretations are represented by the graph. In a way, this objectification of student response actually can point students toward the more mathematical interrogations characteristic of Jockers' macroanalysis. Beyond providing a visual framework for structuring, guiding, and provoking class discussion, sentiment analysis graphs can ultimately generate well-focused topics for written assignments. For example, after discussing the *What Was Mine* graph, students will be well positioned to write on sufficiently narrowed topics about viewpoint, character development, and authorial strategies such as ambiguity, irony, and narrative pace.

Works Cited

Barrett, Lisa Feldman. "Valence Is a Basic Building Block of Emotional Life." *Journal of Research in Personality* 40 (2006) 37, www.elsevier.com/locate.jrp. Accessed 3 Jan. 2020.

Drucker, Johanna. *Graphesis: Visual Forms of Knowledge Production*. Harvard University Press, 2014.

Guerin, Wilfred, *et al*. *A Handbook of Critical Approaches to Literature, 5th ed.* Oxford University Press, 2005.

healthline, "Formication." https://www.healthline.com/health/formication. Accessed 7 Jan. 2020.

Jockers, Matthew L. *Macroanalysis: Digital Methods & Literary History*. University of Illinois Press, 2013.

---. *Text Analysis with R for Students of Literature*. Springer International Publishing, 2014.

Liu, Bing. *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers, 2012.

Moretti, Franco. *Distant Reading*. Verso, 2013.

O'Connor, Flannery. "A Good Man Is Hard to Find." *The Complete Stories*. The Noonday Press, 1999, 117-33.

Rainmaker.FM. *The Writer Files*. "How to Crack the 'Bestseller Code' with Jodie Archer and Matt Jockers: Part Two." November 28, 2017.

Ross, Helen Klein. *What Was Mine*. Gallery Books, 2016.