Original Research Article

Checking Multiple Boxes: Themes Associated With Bicultural Identities

Merrisa Lin[1] and Nairán Ramírez-Esparza[1]

Abstract

Multicultural individuals frequently feel they must check multiple boxes when responding to identity questions. In this investigation, we collected narratives from 402 bicultural individuals about their experiences with multicultural identities and used automated text-analytic tools to examine the valence of the narratives (sentiment analysis) and the most prevalent themes used (Meaning Extraction Method). In addition, we examined the relationships between the narratives and the two dimensions of Bicultural Identity Integration (BII) scales, Harmony and Blendedness. Findings suggested that the narratives were more emotionally positive than negative, and there were six most common themes. Furthermore, themes were significantly associated with BII scales. Our analyses provide deeper insights into the nature of bicultural identity using naturalistic language data.

Keywords

identity, biculturalism, narratives, sentiment analysis, meaning extraction method

My family history is a mosaic of relocations and migrations. So it wasn't strange that one year when I was a kid during the Olympic Games, several athletes on the woman's long-distance race were from countries that my family had connections to. Who are you cheering for, my dad asked with a smile. Well, I answered, I'm shouting for Ireland because we're Irish. I'm shouting for Italy because I was born here. I'm shouting for Jamaica because Mummy was born there. I'm shouting for the U.S. because she grew up there. And I'm shouting for Great Britain because she has a British passport, and that makes me British, too. His smile faded. You are not British, he said in a flat tone. Almost 20 years later, having married a Hungarian, raising kids in Scotland, I've added to the confused mosaic of my family's identity. I've moved on to proudly calling myself a citizen of the world. But when I think back on those years, the desire for an identity, the need for belonging, I still feel uneasy. It's like a kid standing at the edge of the playground when he's not been picked to be on the team, putting on a brave face and saying you know what? I didn't want to play anyway.

—Tomas Sheridan, an interviewee of podcast Rough Translation (Warner et al., 2021b)

[1]University of Connecticut, Storrs, USA

Corresponding Author:
Merrisa Lin, Department of Psychological Sciences, University of Connecticut, Storrs, CT, USA.
Email: merrisa.lin@uconn.edu

I marked off Latino/Latina on a census form. I am first-generation Dominican American. I don't refer to myself as Afro Latina or as a White Latina. I refer to myself as a racially ambiguous woman. And then I got a call from the census. The census was like, excuse me, ma'am. We noticed that you marked off all the boxes. And I was like, yes, as a Dominican, I identify with White and Black and native ancestry. But the census officer persists in trying to get me to choose. They were like, if you had to just mark one box, which box would you mark? And I was like, I wouldn't mark one. They were like, well, if you had a choice, which box would you mark? I'm like, all the boxes.

—Vanessa Lopez, an interviewee of podcast Rough Translation (Warner et al., 2021a)

As globalization expands, it is common to hear stories of people who can check multiple boxes for identity questions. This is particularly common among individuals who identify as bicultural or multicultural, often stemming from life experiences like immigration, having parents from different countries, speaking multiple languages, or being an international traveler (Nguyen & Benet-Martínez, 2007). Negotiating multiple identities can be complicated in many ways, as shown in the quotes above, individuals question their belongingness, and their narratives convey both negative and positive valence. For the past decade, Benet-Martínez and colleagues have led empirical research on Bicultural Identity Integration (BII, for example, Benet-Martínez et al., 2002; Benet-Martínez & Haritatos, 2005; Nguyen & Benet-Martínez, 2007), which examines how bicultural individuals perceive and negotiate their cultural identities (for reviews, see Benet-Martínez et al., 2021; Cheng et al., 2014). BII, conceptualized as perceiving cultural identities as compatible and integrated versus oppositional and separate, has been mostly measured with questionnaires; leaving gaps in understanding the narratives of bicultural individuals. For example, in the quotes above, Thomas discusses cultural belongingness, while Vanessa reflects on her race/ethnic background.

The goal of this investigation was to identify the themes that bicultural individuals use when they feel they must check multiple boxes about their identities. Inspired by the "Rough Translation" podcast, we adapted their question:

Do people ever speak to you like you're a part of a group that you feel you don't belong to? Do you check off one set of boxes in the United States, but identify as something different overseas, or even in your own community?

in this study to examine the experiences of bicultural students living in the United States. We were also interested in understanding how these experiences related to BII scales. For example, in the quotes above, Thomas is questioning his belongingness, but he is also proud of his cultural identities; therefore, he may see his identities as integrated or separated.

BII and Its Psychological Correlates

Biculturalism, or multiculturalism, is a complex and multifaceted aspect of many individuals' lives. According to Benet-Martínez (2012), biculturalism may be defined as meaningful exposure to and experiences with at least two cultural traditions. Bicultural individuals internalize these cultures, and this further impacts how they perceive themselves and the world around them (Huynh et al., 2011). Deriving from Berry's (1990) acculturation strategies for immigrants and ethnic minorities (i.e., assimilation, integration, marginalization, and separation), biculturalism reflects the integration strategy, where individuals identify with both cultures. To understand individual differences in bicultural identity, Benet-Martínez and colleagues (2002) proposed the BII framework. The framework emphasizes bicultural individuals' subjective perceptions of how much their cultural identities intersect or overlap. BII includes two distinct but related dimensions—cultural

harmony versus conflict and cultural blendedness versus compartmentalization (hereafter harmony and blendedness; Benet-Martínez & Haritatos, 2005; Cheng et al., 2014; Huynh et al., 2011, 2018). Harmony captures the degree of harmony versus tension felt between the two cultural orientations. In other words, the extent to which a person feels their two cultures exist compatibly within the self as compared to feeling conflicted. Blendedness highlights the extent to which individuals view themselves as part of a blended cultural identity as opposed to seeing their two cultures as separated from one another. This dimension captures how much bicultural individuals perceive overlap versus distance between their two cultural orientations (Benet-Martínez & Haritatos, 2005; Cheng et al., 2014; Huynh et al., 2011, 2018).

Harmony and blendedness are associated with different sets of psychological constructs. The Harmony dimension captures the affective component of managing two cultures, while the Blendedness dimension captures the cognitive and behavioral aspects of bicultural experiences (Huynh et al., 2018). Harmony is positively associated with personality traits of emotional stability, lower psychological distress, greater well-being, stronger feelings of belonging, and positive affect toward their cultural groups (e.g., Huynh et al., 2018; for discussions of related psychological constructs, see Benet-Martínez & Haritatos, 2005; Huynh et al., 2011). Harmony is influenced by contextual strains, with greater harmony related to less perceived discrimination and more tolerance toward different cultural values (Leong & Ward, 2000). On the other hand, the Blendedness dimension is linked to cognitive and performance-related challenges of BII. This dimension is positively associated with the personality traits of openness to experience, bicultural competence, and linguistic fluency. Individuals high in blendedness have more exposure to their different cultures (e.g., live in a culturally diverse environment), thus they may develop higher proficiency in related languages, which facilitates the formation of a combined identity (e.g., Benet-Martínez & Haritatos, 2005; Huynh et al., 2011, 2018). For instance, Chinese American biculturals with lower blendedness have greater language barriers, live in more culturally isolated surroundings, and are less open to new experiences (Benet-Martínez & Haritatos, 2005).

All in all, Harmony and Blendedness are two independent but related dimensions of BII. Whereas harmony represents the feelings and attitudes toward BII, cultural blendedness indexes the cognitive and behavioral organizations of the two cultural orientations. The BII framework emerges from the differences in both harmony and blendedness, with individuals exhibiting any combination of high or low harmony and blendedness (Benet-Martínez, 2012). For example, a Taiwanese American bicultural may feel conflicted between their two cultures' values but practice both cultures' traditions.

Multicultural Identity and Narrative Methodologies

Bicultural and multicultural individuals construct their bicultural identity through their life experiences. However, there is still much to be understood about the content and meaning of these experiences, especially with respect to how bicultural individuals experience their identities. The narrative approach offers a way of exploring how individuals draw from their unique experiences to understand their identities. However, research in this area is limited, and the methodologies used are diverse. For example, Phinney and Devich-Navarro (1997) qualitatively analyzed interviews of African American ($N = 52$) and Mexican American ($N = 46$) biculturals while Yampolsky et al. (2013) assessed narrative coherence from interviews with a diverse sample of Canadian biculturals ($N = 22$). Syed and Azmitia (2010) coded narratives from participants of diverse ethnic backgrounds ($N = 70$) to extract four identity narrative themes (awareness of difference, awareness of underrepresentation, experience of prejudice, and positive connection to culture/ethnicity).

Lilgendahl and colleagues (2018) asked a diverse sample of bicultural students (Study 1, $N = 77$) and bicultural adults (Study 2, $N = 50$) to share memories that triggered heightened awareness of being bicultural (e.g., "Think of a specific memory during which you became highly aware of your bicultural identity …"). The participants also completed the BII Scale Version 2 (Huynh et al., 2018). The stories were then coded by hand based on three factors: exploratory processing (how much the narrator uses themes about exploring what it means to be bicultural), the emotional valence of the past event, and the emotional valence of the ending (positive and/or negative). The authors found that what matters for BII was the ending of the story, not the narrative valence of past events. Participants who ended their stories with a positive resolution showed higher scores on harmony and blendedness.

An individual's bicultural identity is constructed through the narration of a variety of different memories that reflect their bicultural experiences. There is still much to learn about how multiple identity experiences influence bicultural individuals' BII development. For instance, Lilgendahl et al. (2018) found that each participant's narrative of their bicultural identities was unique. Therefore, our study adds to what is known about bicultural experiences by examining the themes that bicultural individuals talk about when reflecting on their identities and how these themes relate to BII.

In this study, we build on Lilgendahl and colleagues' (2018) work by conceptualizing bicultural identity as a part of people's everyday experiences that shapes an individual's self-understanding. We used a generic prompt that broadly captures an individual's multiple identity experiences and focuses on the overall emotional valence to explore the psychological processes in bicultural identity narratives. Our approach uniquely combines top-down and bottom-up methods. We use text-analytic tools to automatically detect the valence of the narratives (i.e., sentiment analysis, Hutto & Gilbert, 2014; Li & Hovy, 2017) and the most common themes used in the narratives (i.e., the Meaning Extraction Method [MEM]; Boyd, 2017; Chung & Pennebaker, 2008a, 2008b). These approaches have some advantages over the previous studies discussed. First, the methods avoid imposing themes predetermined by the researchers and rely on the most common content words used in the narratives. Furthermore, these methods permit the inclusion of a wide range of themes without focusing on only a few pre-established themes. Finally, since we use advanced automatic text-analytic tools, we can analyze narratives from a larger number of individuals, given that we do not rely on manual coding done by judges.

Sentiment Analysis and the MEM

To understand the valence of the narratives, we adopted the sentiment analysis technique, which is a top-down process that identifies the valence of words that reflect authors' attitudes in a given text corpus (i.e., positive, negative, and neutral; Li & Hovy, 2017). This approach relies on the sentiment lexicon, which is a list of words that are labeled according to their semantic orientation as positive (e.g., *love*, *enjoy*, and *happy*), negative (e.g., *ugly*, *sad*, and *bad*), or neutral (context-dependent, for example, "catch" a fish is considered neutral) classes (Liu, 2010). The sentiment lexicon has been developed and validated over decades by experts in psychology, linguistics, and sociology. For instance, a widely used sentiment lexicon, Linguistic Inquiry and Word Count (LIWC) dictionary, contains nearly 1,000 semantic orientation-based words (Pennebaker et al., 2022). A common model in this technique is the bag-of-words. The model creates a list of unique words based on the text corpus and compares the words in each text entry with the list. The model quantifies words by computing whether a given word exists and its frequency in each text entry (Y. Zhang et al., 2010). Through incorporating the emotion-related words in the predetermined dictionary and converting words into numbers, the sentiment analysis computes the valence of words in a given text corpus.

Next, to examine what themes bicultural individuals talk about when narrating their identities, we employed the MEM (Boyd, 2017; Chung & Pennebaker, 2008a, 2008b). The MEM is

a bottom-up topic modeling technique that extracts psychologically meaningful themes from natural language data and is considered a mixed-method approach (Boyd, 2017; Chung & Pennebaker, 2008a, 2008b). The method uses principal component analysis to identify clusters of words that frequently co-occur across the text corpus. These clusters are then labeled using an inductive approach to identify the most common themes. The assumption that different words belonging to an overarching theme will tend to be used together provides the foundation to draw inferences about what is being discussed and in what quantities, which can be applied to any topic in any language (Boyd & Pennebaker, 2016). The MEM has been used in the past to study self-schema across cultures. Ramírez-Esparza and colleagues (2012) used the MEM to examine open-ended personality descriptions, identifying the most salient personality dimensions in American and Mexican cultures. Building on this work, Rodríguez-Arauz et al. (2017) studied whether Mexican American bilinguals' self-schemas change when alternating between English and Spanish. The straightforward steps needed to do MEM make it possible to easily apply, interpret, and modify a large body of text, which avoids the problems that come with more complex topic modeling methods.

In sum, both sentiment analysis and the MEM have been valuable tools for understanding sociocultural and psychological concepts. Therefore, the present study uses these techniques to examine the themes that bicultural individuals living in the United States use when talking about their experiences with multiple identities. Through integrating sentiment analysis and MEM, this investigation also sheds light on new approaches that can be used to study bicultural identities.

Current Research

In this study, we adapted the question from the Rough Translation podcast to prompt bicultural individuals to reflect on their multicultural experiences. Importantly, we did not exclusively prompt participants to think about their bicultural experiences, as in Lilgendahl et al. (2018). Our overarching goal is to explore the topics that emerge when bicultural individuals narrate their multiple identities in general. This subtle shift in the prompt acknowledges the intersection of bicultural identity with various other identities, mirroring real-world dynamics. In addition, participants also responded to the BII Scale Version 2 (Huynh et al., 2018) to identify the degree to which they see their identities as harmonious and/or blended. In all, we aim to integrate quantitative scale measures with text-based narratives to understand how BII and identity narratives interplay with each other. While we do not make specific predictions since our analyses are exploratory, we are guided by previous studies to formulate general expectations.

Our first goal was to use sentiment analysis to capture the overall valence of the narratives. We aim to complement Lilgendahl et al.'s manual coding of bicultural individuals' memories, in which they found that narratives were somewhat negative, although not significantly related to BII. Our computer-based automatic methodology will reduce potential bias involving human coders.

Second, we adopted the MEM to identify themes in bicultural individuals' narratives about their multiple identities. Phinney and Devich-Navarro (1997) found that Mexican American and African American biculturals based their sense of being bicultural on experiencing uncertainty about their cultures, identifying with their ethnic culture, identifying with the American culture, feelings of inclusion in American society, understanding cultural roots, being bicultural in physical terms, and being bilingual. Also using a qualitative approach, Huynh et al. (2018) found that bicultural individuals mentioned the compatibility and similarity of their cultures, the uncertainty and problems of being bicultural, and the benefits of being bicultural (Huynh et al., 2018). Accordingly, we speculate that the current bicultural sample may narrate their experiences using similar themes.

Third, we observe the association between the valence of the narratives and BII. While Lilgendahl et al. (2018) reported that positive tones in narrative endings correlated positively

with both harmony and blendedness, the overall valence of the text was unrelated to BII. Our study aims to explore whether our different methodologies yield similar results.

Finally, we observe the association between the themes and BII. Phinney and Devich-Navarro's (1997) study gives us a hint that themes may relate to BII in different ways. For example, the authors identified blended and alternating biculturals. The authors found that blended biculturals expressed good feelings of being bicultural and practiced both cultures equally, whereas alternating biculturals indicated closeness with their ethnic-cultural group and mentioned their physical traits and bilingual experiences. Accordingly, these findings suggest that different levels of BII may be associated with different bicultural experiences.

Method

Participants

A total of 910 participants (mean age = 19.68, *SD* = 3; 703 females, 205 males, and two others) participated in the current study. However, this study aimed to examine people with bicultural backgrounds; thus, monocultural participants were excluded. Bicultural identification was left open to report on the part of the participants; however, we framed bicultural individuals as people who identify with more than one culture, one being American and the other being any other cultural identification. The final analytic sample included 402 bicultural participants (mean age = 19.39, *SD* = 2.3; 300 females, 102 males). Participants were recruited from the University of Connecticut (*n* = 345, 85.8%), the University of Florida, and Texas Tech University (*n* = 57, 14.2%). Most students participated in the study to earn course credits for the introductory psychology classes. Some participants earned extra course credits for their participation upon agreement with their course instructors. Participants were also recruited through word-of-mouth and advertisements within student organizations. Because there was no differential effect of the method of recruitment, we collapsed participants recruited from the University of Connecticut, the University of Florida, and Texas Tech University.

Most of the participants were undergraduates, except for 10 (2.5%) who indicated that they had postgraduate or equivalent degrees. When asked about their parents' socioeconomic status (SES) on a 5-point Likert-type scale ranging from 1 (*working class*) to 5 (*upper class*), participants indicated an average of 2.88 (*SD* = 1.03). Regarding participants' cultural backgrounds, the majority of the participants identified as affiliating with Latinx cultures (*n* = 182, 45.3%), while others were affiliated with European American cultures (*n* = 74, 18.4%), Chinese cultures (*n* = 59, 14.7%), South Asian cultures (*n* = 48, 11.9%), African-related cultures (*n* = 17, 4.2%), Brazilian cultures (*n* = 9, 2.2%), other East Asian cultures (*n* = 8, 2%), and Middle Eastern cultures (*n* = 5, 1.2%). Regarding participants' racial backgrounds, the majority identified as White (*n* = 207, 51.5%), followed by Asian (*n* = 111, 27.6%), and Black or African American (*n* = 25, 6.2%). Other participants were grouped as other races (e.g., Native American or Alaska Native, Native Hawaiian or Other Pacific Islander, more than one race, and others; *n* = 19, 4.6%). Most participants were born in the United States (*n* = 306, 76.1%), while 96 participants (23.9%) were born outside of the United States.

General Procedure

The study was conducted in an online questionnaire format via the software Qualtrics. An online link was provided to participants upon signing up for the study. Participants first read the information sheet and the study instructions. Upon giving consent, participants were instructed to complete a set of demographic questions. These questions included age, gender, SES, education background, race, ethnicity, language background, place of birth, and the length of residence in

the United States. Culture-related questions, such as cultural affiliations and heritage cultures, were also included. Then, participants were directed to the bicultural identity measure. Finally, participants were prompted with an open-ended identity question and described their identity for 5 minutes. Upon completion, a debriefing sheet was presented. Current research is part of a project investigating humans' psychological processes. Measures that are out of the scope of the present study are not discussed. This study received approval from the University of Connecticut Institutional Review Board (IRB). All participants followed the same procedures according to the IRB protocols to complete the study.

Measures

Bicultural Identity Integration Measure. We used the Bicultural Identity Integration Scale Version 2 (BIIS-2) to measure the extent to which participants integrate their multiple cultural identities (Huynh et al., 2018). BIIS-2 consisted of 17 items, and participants rated their level of agreement on a 5-point Likert-type scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). In this study, participants were asked to complete the measure in terms of how they feel about American culture and their other primary cultural identification (e.g., heritage culture). The scale contains two separate dimensions—harmony versus conflict and blendedness versus compartmentalization. The Harmony subscale contains 10 items and relates to the values, attitudes, and affections of being bicultural. The subscale measures the extent to which an individual feels harmonized or conflicted between their cultural identities. Sample items for the Harmony subscale are "I feel caught between my other culture(s) and my American culture" and "I find it easy to harmonize my other culture(s) and my American culture." Cronbach's alpha for the Harmony subscale was $\alpha = .86$. The second dimension, the Blendedness subscale, contains seven items. The subscale measures the extent to which individuals feel their different cultural identities combine to create a blended cultural identification in contrast to experiencing the culture as separated. Sample items for the Blendedness subscale are "I feel part of a combined culture" and "I cannot ignore the other culture or American side of me." Cronbach's alpha for the Blendedness subscale was $\alpha = .76$. The two subscales were significantly correlated with one another ($r = .11$, $p = .03$). This corresponded to previous studies that identified these two subscales as being weakly positively correlated as they are related but theoretically independent concepts (e.g., Benet-Martínez & Haritatos, 2005; Cheng et al., 2014; Huynh et al., 2011, 2018; Lilgendahl et al., 2018).

Participants with higher scores on both dimensions tend to better develop compatible bicultural identities, whereas those with lower scores have more difficulty integrating both cultures into their identities. For additional information, we also included a composite version, referred to as "total BII," calculated by averaging all 17 items together ($\alpha = .81$).

Narrative Prompt. To identify bicultural individuals' experiences with their identities, we adapted the question from the Rough Translation podcast (Warner et al., 2021b). Specifically, the participants read the following:

> Do people ever speak to you like you are a part of a group that you feel you don't belong to? Do you check off one set of boxes for identity questions (e.g., race, ethnicity, gender, religion etc.), but identify as something different overseas, or even in your own community? For the next 5 minutes, please think about the above questions and share your experiences.

This prompt diverges from an exclusive focus on bicultural identities and instead centers on the broader spectrum of experiences associated with having multiple identities. This is different from previous studies that instructed participants to reflect on their bicultural experiences with heightened emotions (e.g., Lilgendahl et al., 2018; Syed & Azmitia, 2010). Our goal is

to understand the overarching themes participants discuss when thinking about their multiple identities. Accordingly, we prompted participants to recall instances where they felt they must check multiple boxes to facilitate narratives about their experiences with multiple identities.

A 5-min countdown timer was presented along with a large blank text field for typing. The timer was used to ensure all participants had the same amount of time to answer the question. Participants were instructed to write for the full 5 minutes; however, participants could continue writing if they wished. Furthermore, this approach allowed us to reduce mental exhaustion while making sure the duration was long enough for participants to genuinely narrate their multiple identity experiences.

Text Analytical Strategies

Data Preprocessing and Preparation. The responses to the identity measure were cleaned using the *Tidyverse* package in the statistical software R (R Core Team, 2020; Wickham et al., 2019). Empty responses and texts containing fewer than 50 characters—which is about 10 words or one full sentence—were omitted. We removed these responses as they suggested that participants did not take the 5-min writing task seriously. The average word count for the responses was 79.65 (*SD* = 48.04), with a minimum of 10 words and a maximum of 411 words, showing that participants composed a few sentences into a short paragraph. Past studies using the MEM approach to study self-schema revealed that bilingual participants wrote 186.3 words on average for the 15-min time frame (Rodríguez-Arauz et al., 2017). Given that we adopted a 5-min time frame, our data provided sufficient length with meaningful information about participants' experiences. Preprocessing resulted in 373 responses being retained. This sample was included in current study analyses.

Analyzing Narrative Valence—Sentiment Analysis. To explore the valence of bicultural individuals' narratives, we performed sentiment analysis on participants' open-ended responses. We used the freely accessible package *Vader* in the statistical software R (R Core Team, 2020; Roehrick, 2020). Vader analysis relies on a predetermined dictionary that maps text corpus features to emotion polarities and intensities, and then generates the sentiment scores accordingly. It is a simple, human-centric, and computationally efficient approach that has been empirically validated. Vader performed as well as, in most cases, better than other highly regarded sentiment analysis tools such as LIWC (Pennebaker et al., 2022), ANEW (Bradley & Lang, 1999), and General Inquirer (Stone et al., 1966). VADER keeps the advantages of traditional sentiment lexicons, like having complete and easy-to-use dictionaries. It also takes into account emoticons, acronyms, initialisms, and slangs, which are all words that can show emotion. In addition, it has received approval from human raters and narrative-based social media texts and blogs (Hutto & Gilbert, 2014). In our analysis, a compound score is generated for each text corpus. The compound score is a metric that calculates the sum of the valence scores of each word in the text, weights according to rules such as punctuation amplifiers and negation polarity switches, and then normalizes between −1 (*most negative valence*) and +1 (*most positive valence*). In other words, compound scores are not the average counts of positive, negative, and neutral words that cancel out contributions in opposite directions. The scores account for rule-based enhancements to provide objective sentiment information, such that higher compound scores indicate more positive contents. For example, words like *love*, *enjoy*, and *happy* convey positive sentiment (Hutto & Gilbert, 2014).

Extracting Narrative Themes—MEM. To extract themes from the narratives, we used the software Meaning Extraction Helper (MEH, Boyd, 2018), which automatically extracts the most meaningful themes according to the MEM (Boyd, 2017; Chung & Pennebaker, 2008a, 2008b). The MEH uses word-based text-analytic tools to identify the most commonly used content words in a text

corpus. The MEH software converts natural language data into statistical measures that can be analyzed quantitatively. That is, the single words in the text corpus are converted into frequency statistics. The MEH also allows adding extra conversions and stop words that are not yet included in the program. Conversions refer to changing the words with the same roots into one word, changing common abbreviations into open forms, and correcting misspellings. Stop words are words that appear in text but have little to no meaning. These words are excluded from the MEH analysis (Boyd, 2018). In the current study, we excluded words such as *box*, *time*, *person, and people and* converted words that share the same root or meaning (e.g., grandmother and grandfather were converted to grandparent). Many of the front-end tasks of the MEM are automated by the MEH. These tasks are also used by other topic modeling methods and are easy to use with advanced language analysis methods like Latent Dirichlet Allocation and open vocabulary approaches. These benefits enhance the replicability and reproducibility of research when employing alternative topic modeling techniques (Boyd & Pennebaker, 2016). Words that were observed in at least 1.2% of the responses were included in the analyses (see Boyd, 2018, for considerations of observation cutoffs). Then, the binary (also called "one hot encoding") data were used in a principal component analysis with varimax rotation. Factors were extracted based on a Scree plot of Eigenvalues for the principal components. Factors at the elbow bend that had Eigenvalues above 1 and that added a significant increase to the cumulative percent variance were selected (Cattell, 1966).

Results

To assure that participants were attentive to the narrative prompt, we conducted an attention check midway through the study (i.e., 41 questions following the narrative prompt). Although we expected a reasonable percentage of incorrect answers given the timing of the check in the study, 87.9% of participants passed the attention check, indicating attentiveness. Moreover, the study was only accessible to students at the University of Connecticut, the University of Florida, and Texas Tech University, despite being administered online. To prevent external influences, participants were required to verify their identity and remain on the study page. The first author looked through the data and confirmed that each response had its own content, sentence structure, and grammar. This ruled out the possibility that AI was responsible for writing the responses. In addition, our data collection (from January 2021 to June 2022) predates the popularity of language model-based chatbots like ChatGPT that were launched in November 2022, further affirming the authenticity of participant responses.

The below sections are divided into three parts: First, we present the results of the sentiment analysis. Second, we describe the themes that resulted from the MEM. Finally, we test the inter-correlation between the valence of the narratives, themes, and BII scales.

Narrative Valence

The content and descriptive statistics of the valence of the narratives are presented in Table 1. The average compound valence score was 0.31 (*SD* = 0.46) in the −1 (*most negative*) and +1 (*most positive*) ranges. The results demonstrated that participants have relatively positive sentiments when narrating their bicultural identity experiences.

Narrative Themes

From the MEM, five factors were extracted, which together accounted for 8.46% of the total variance. Compared with scale-level quantitative data, natural language data is sparser in terms of word co-occurrences. People are not forced to select a score to represent their responses; instead, they

Table 1. Content and Descriptive Statistics of Study Variables.

Variables	Example words and items	M (SD)	Minimum	Maximum
Sentiment analysis				
Narrative valence	positive: love, nice, good; negative: sad, bad, hurt	0.31 (0.46)	-0.86	0.99
Narrative themes				
Cultural belongingness	outsider, immigrant, root, grow, born	2.05 (2.47)	0	30
Mixed cultural experience	mix, culture, connect, friend, household	1.36 (1.96)	0	29
Experience of prejudice	gender, race, religion, discriminate, mark	0.82 (1.51)	0	23
Physical characteristics	skin, look, hair, pale, dark	1.27 (1.62)	0	13
Language experience	English, Spanish, accent, language, speaker	1.19 (1.61)	0	13
Race/ethnic issues	race, ethnic, white, black, embrace	1.3 (1.32)	0	9
Bicultural identity integration				
Total BII	(combination of two subscales below)	3.37 (0.58)	1.88	4.88
BII Harmony	I find it easy to harmonize my two cultures	3.18 (0.81)	1.3	5.0
BII Blendedness	I feel part of a combined culture	3.63 (0.67)	1.0	5.0

Note. For the narrative themes, the mean represents the mean frequency score of the given theme. This value demonstrates the relative degree to which participants invoked or attended to a given theme. Any given narrative may contain multiple themes. Indeed, the mean number of themes present in any given narrative was 3.71 (median = 4; SD = 1.51), highlighting that the majority of participants mentioned several themes.

freely express themselves using text responses. In other words, language exhibits significantly greater variability in the multiplicity of ways it can be expressed compared to scale or Likert-type data. Thus, lower explained variance and loadings are expected (e.g., Chung & Pennebaker, 2008a; Rodríguez-Arauz et al., 2017). Past research using the MEM to study a variety of texts (e.g., X, sexual self-schema, self-concepts) across populations has recorded explanatory variance ranging from 8.79% to 14.7% (Chung & Pennebaker, 2008a; Ikizer et al., 2019; Rodríguez-Arauz et al., 2017), indicating that the explanatory variance in the current work is acceptable.

As can be seen in Table 2, four of the five factors brought together a group of content words that are psychologically meaningful and coherent. However, there were almost equal numbers of positive and negative factor loadings in Factor 5. This meant that it had two separate components, which were named Factor 5Pos and Factor 5Neg. To better understand the factors, we computed the frequency scores for the four factors that were consistent in positive loadings and the two subfactors (i.e., Factor 5Pos and Factor 5Neg). These six factors are interpreted as the six narrative themes for the current study. For each theme, we summed the occurrence of content words that belong to the given theme using the binary data. The new frequency scores quantified each theme in individual responses and provided non-negative scores, as the minimum frequency of a content word is zero. The frequency scores facilitated interpretation, as higher scores indicated that participants are invoked or attended to a given theme. Thus, we used the frequency scores for further analyses. Finally, we used a qualitative approach to label the six extracted themes. The labeling was done based on the meaning of the content words as well as reading the narratives with the highest scores in each of the factors. Table 1 shows the content and descriptive statistics for each MEM theme. The Supplemental Materials show illustrations of the six themes based on

Table 2. Narrative Themes: A Varimax-Rotated Principal Component Analysis.

Narrative themes of bicultural individuals' multiple identity experiences

Factor 1: Cultural belongingness		Factor 2: Mixed cultural experience		Factor 3: Experience of prejudice		Factor 4: Physical characteristics		Factor 5Pos: Language experience		Factor 5Neg: Race/ethnic issues	
life	.45	culture	.27	aspect	.49	family	.23	spoke	.26	embrace	-.27
weird	.42	friend	.21	activity	.49	complete	.21	hear	.23	category	-.22
involve	.41	household	.50	adopt	.46	dark	.63	realize	.31	race	-.31
family	.41	play	.48	mark	.44	hair	.62	learn	.28	white	-.31
show	.40	quick	.46	extreme	.43	skin	.62	english	.37	issue	-.30
root	.39	best	.45	explain	.41	pale	.48	spanish	.36	racial	-.29
spoke	.38	embrace	.44	strange	.41	told	.44	talk	.32	group	-.25
aware	.37	type	.43	work	.40	tone	.39	language	.30	identity	-.24
member	.36	student	.42	mean	.40	light	.38	travel	.29	black	-.24
culture	.33	past	.41	gender	.40	half	.37	grade	.29	easy	-.24
friend	.33	tradition	.39	right	.39	parent	.32	speaker	.27	bit	-.22
young	.32	nationality	.38	form	.38	color	.32	accent	.27	privilege	-.21
comfortable	.32	difficult	.34	child	.38	grandparent	.27	school	.24	ethnic	-.21
move	.32	realize	.33	term	.37	honest	.27	kid	.24	categorize	-.20
point	.32	learn	.33	religion	.37	side	.24	native	.24		
sense	.31	fall	.33	personally	.36	society	.23	fun	.23		
grow	.31	portuguese	.32	woman	.34	pass	.22	start	.22		
said	.30	college	.31	discriminate	.31	look	.21	lost	.22		
relative	.30	mix	.31	happen	.31	white	.29				
immigrant	.29	year	.31	assume	.30						
outsider	.28	category	.30	man	.27						
born	.27	americanize	.29	asian	.25						
allow	.27	complete	.27	race	.28						
strongly	.27	raise	.26								
love	.26	connect	.25								
hear	.24	place	.25								
accept	.22	ethnicity	.24								

Table 2. (continued)

Narrative themes of bicultural individuals' multiple identity experiences

Factor 1: Cultural belongingness		Factor 2: Mixed cultural experience		Factor 3: Experience of prejudice	Factor 4: Physical characteristics	Factor 5Pos: Language experience	Factor 5Neg: Race/ethnic issues
live	.22	activity	.39				
call	.20	right	.36				
change	.20						
heritage	.20						
see	.20						
connect	.23						
explain	.27						
parent	.27						
accent	.22						

Note. The factor loadings of content words are reported. Five factors were originally extracted from the varimax-rotated principal component analysis. Factor 5 was then separated into Factor 5Pos and Factor 5Neg according to the content words' factor loading directionality.

Table 3. Sample Responses of the Narrative Themes.

Themes	Sample responses
Cultural belongingness	"Most of my family is Jamaican, so I grew up associating with being Jamaican. Especially since both of my parents are Jamaican. However I remember in middle and high school I had a Jamaican friend who was also born in Jamaica, but she spent more time there then I did (I moved to the U.S. when I was 1). At some points when I was talking to her I felt like I wasn't "Jamaican enough" because I only spent one year there and she knew more about Jamaica. She moved to the U.S. around age 8 or 10 and she visited Jamaica a couple of times since she moved and I still haven't gone back yet."
Mixed cultural experience	"I do not think that people speak to me like I am part of a group that I do not belong to. I identify as American and Puerto Rican. So, being placed into one or another does not bother me. Being raised in a Puerto Rican influenced household, allowed me to relate/identify with the Puerto Rican culture. Additionally, being raised in American, especially in American schools (majority of students being white Americans) allows me to understand and identify with other White Americans as well."
Experience of prejudice	"People assume things based on my race and gender. I identify as an Asian American and people assume that I should speak Mandarin. As well as trying to explain my family dynamic to some people so why I don't look like my parents or siblings and then trying to simplify what it means to be adopted. As a female we have come such a long way in terms of women's rights activism and yet we are still so far from equal to men. We can be paid less, we might be assumed to not be as able in some aspects but more able in others. As for my religion, I am agnostic meaning that I am not as extreme as atheists who believe there is no such thing as gods. I mark that I am non-religious or agnostic on forms because although I was baptized and my family follows Christianity, I don't personally identify with that group."
Physical characteristics	"A lot of people think that I am not Mexican because of my coloring. I have light skin with light eyes and brunette/blonde hair. I get told that I am not Mexican, and if my family is I must be adopted. Also because of my economic status, I often got assimilated with the stereotype of the Mexican culture."
Language experience	"As a Spanish native speaker I often felt out of place speaking English when I was starting to learn it, some kids would make fun of my accent and my wrong pronunciation and I would hate reading in class. When coming to the U.S I realized that regardless of where you are from, if you speak Spanish you are welcomed into the Hispanic community, even though outside the U.S many Spanish speaking countries don't like each other."
Race/ethnic issues	"People speak to me as if I belong to certain groups which I don't necessarily believe myself to be. Being Dominican and Colombian growing up in New York, primarily White people believe me to be ghetto and "Black". Black people don't fully accept me as one of them. I am in between both categories and I have trouble associating with either of them. When filling out race boxes for my identity I feel like I don't belong in any of the categories. I am not White nor Asian. The only race that I can identify with is Black but even that goes back many generations and I feel like it doesn't properly describe who I am today. I identify as Hispanic/Latino."

how often certain words were used (Supplemental Figure S1) and the frequency distribution of each theme (Supplemental Figure S2).

Theme 1: Cultural Belongingness. Sample responses are presented in Table 3. The first theme included words such as *culture*, *family, life*, *move*, *born*, and *grow*. These words related to the experiences participants had throughout their lives. Words such as *outsider, weird, immigrant,*

accept, and *root* were also included in this theme, showing that participants talked about their experiences of moving between cultures and trying to find belongingness. The sample response demonstrated that the participant mentioned their experiences in the United States and their belongingness related to being both an American and a Jamaican. The participant also mentioned how they experience both cultures while feeling not enough for either one. The mean frequency score of cultural belongingness theme was 2.05 ($SD = 2.47$).

Theme 2: Mixed Cultural Experience. The second theme included words such as *culture*, *tradition*, *embrace*, *complete*, *connect*, and *mix*. These words center around accepting and embracing multiple identities. This theme also included words such as *household*, *student*, *friend*, and *college*, which captured participants' mixed cultural experiences in different contexts. The sample response showed that the participant identified with both American and Puerto Rican cultures. They experienced American cultures in school and practiced Puerto Rican cultures at home. The mean frequency score of this theme was 1.36 ($SD = 1.96$). See Table 3 for sample responses.

Theme 3: Experience of Prejudice. Words such as *mark*, *assume*, *explain*, *discriminate, gender*, *religion*, *race*, *woman*, and *man* were included in the third theme. These words were relevant to participants' experiences being labeled as being part of a social group that they may or may not identify with. The theme also captured how others interact with the participants based on these social labels in a prejudiced way. For instance, the sample response showed that the participant mentioned their identification as Asian American, a woman, and an agnostic. Because the participant is Asian American, people assume that they speak Mandarin, but the participant does not. For this theme, the mean frequency score was 0.82 ($SD = 1.51$). See Table 3 for sample responses.

Theme 4: Physical Characteristics. The fourth theme related to the experiences participants had based on how they looked. This theme included physical traits related to words such as *hair*, *skin*, *look*, *color*, *tone, dark*, *pale*, *light*, and *white*. In this theme, participants mentioned their experiences related to their physical characteristics, such as skin and hair color. For instance, the sample response showed that the participant's lighter skin, bluer eyes, and blonde hair did not reveal their Mexican background. The mean frequency score of the physical characteristics theme was 1.27 ($SD = 1.62$). See Table 3 for more details.

Theme 5: Language Experience. This theme included words such as *spoke*, *speaker*, *learn*, *talk*, and *language*. These words captured the experience of speaking different languages. Words such as *accent*, *native*, *English*, and *Spanish* were also included in this theme, showing that accent and nativeness in languages such as English and Spanish played some roles in individuals' identities. The sample response presented the participant's experiences of learning the host language (i.e., English) and their feelings when using their heritage language (i.e., Spanish). The mean frequency score of this theme was 1.19 ($SD = 1.61$). See Table 3 for more details.

Theme 6: Race/Ethnic Issues. This theme captured responses specific to racially or ethnically related experiences. Different from the physical appearance theme, this theme especially focused on the influences of race and ethnicity on participants' identities. This theme included words such as *category*, *categorize*, *group*, *identity*, *race*, *racial*, *ethnic*, *white*, and *black*. These words captured the racial categories participants identify with or belong to. Words such as *embrace*, *issue*, *easy*, and *privilege* were also included in this theme, demonstrating the experiences participants had because of their racial and ethnic backgrounds. For instance, the sample response showed that people do not categorize the participant as the racial group they identify with, nor do the participants themselves feel a sense of belonging in racial categories on the standardized forms. The mean frequency score of race/ethnic issues theme was 1.3 ($SD = 1.32$). See Table 3 for sample responses.

Table 4. Correlations of Narrative Valence, Themes, and Bicultural Identity Integration.

Variables	1	2	3	4	5	6	7	8	9	10
Sentiment analysis										
1. Narrative valence										
Narrative themes										
2. Cultural belongingness	.25**									
3. Mixed cultural experience	.21**	.38**								
4. Experience of prejudice	.06	.07	.09							
5. Physical characteristics	-.06	.37**	.19**	.12*						
6. Language experience	.10	.32**	.19**	-.03	.06					
7. Race/ethnic issues	-.05	-.07	.11*	.23**	.17**	-.24**				
Bicultural identity integration										
8. Total BII	.04	-.12*	-.04	.07	-.13**	-.13**	.06			
9. BII Harmony	.02	-.22**	-.13*	.03	-.15**	-.17**	.04	.88**		
10. BII Blendedness	.05	.11*	.13*	.09	-.02	.02	.05	.57**	.11*	

Note. For the narrative valence, higher score indicates a more positive valence.
*$p < .05$. **$p < .01$.

Correlations for Narrative Valence, Themes, and BII

The content and descriptive statistics of the total BII, Harmony, and Blendedness subscale are presented in Table 1. Table 4 shows the intercorrelations among the variables of interest: the valence of the narratives, the six narrative themes, and BII (total score, Harmony subscale, and Blendedness subscale). To explore who is more or less likely to narrate a given theme, we also performed analyses on demographic variables (e.g., age, gender, SES, ethnicity, race). See Supplemental Materials for details.

Narrative Valence. Our results showed that the narrative valence positively significantly correlated with cultural belongingness ($r = .25, p < .01$) and mixed cultural experience ($r = .21, p < .01$) themes. The more positively toned participants were in their identity narratives, the more frequently they mentioned cultural belongingness and mixed cultural experience. In regard to the correlations between the narrative valence and BII (total score, Harmony subscale, and Blendedness subscale), no significant correlations were found. See Supplemental Materials for additional analyses on the correlation between the ending narrative and BII.

Narrative Themes. Our analyses showed that cultural belongingness ($r = -.12, p = .02$), physical characteristics ($r = -.13, p = .01$), and language experience ($r = -.13, p = .01$) themes negatively significantly correlated with BII total score. Next, we found that cultural belongingness ($r = -.22, p < .01$), mixed cultural experience ($r = -.13, p = .01$), physical characteristics ($r = -0.15, p < .01$), and language experience ($r = -.17, p < .01$) themes negatively significantly correlated with Harmony subscale. Cultural belongingness ($r = .11, p = .04$) and mixed cultural experience ($r = .13, p = .01$) themes positively significantly correlated with Blendedness subscale. The experience of prejudice and race/ethnic issues themes did not correlate significantly with BII total score, Harmony subscale, or Blendedness subscale. The intercorrelations between narrative themes are reported in the Supplemental Materials.

In summary, the pattern of correlations above suggests: (a) bicultural individuals who discuss less about cultural belongingness, physical characteristics, and language experience are better at integrating their bicultural identities; (b) bicultural individuals who talk less about cultural belongingness, mixed cultural experience, physical characteristics, and language experience are

better at harmonizing their multiple cultural identities; and (c) bicultural individuals who talk less about cultural belongingness and mixed cultural experience are better at blending their multiple cultural identities. All in all, these results suggest that when narrating identities, bicultural individuals emphasize different themes, and the frequency of mentioning each theme relates to the extent to which they integrate their bicultural identities.

Discussion

With globalization, navigating multiple cultural streams is becoming increasingly prominent, highlighting the importance of understanding bicultural individuals' experiences for identity development (Meca et al., 2019). Our overall goal was to explore the experiences of U.S. bicultural students who feel the need to check multiple boxes when answering identity questions.

Narrative Valence

The first aim of the present study was to examine the narrative valence when bicultural individuals share their bicultural experiences, which we found to have a relatively positive valence. This differs from Lilgendahl et al.'s (2018) findings, where bicultural narratives tended to be somewhat negative. This discrepancy may stem from differences in the prompts used. Lilgendahl et al. framed BII as an identity challenge in which participants were prompted to recall a detailed account of incidents where they were highly aware of their bicultural identity and its impacts. This approach is typically used in the narrative identity literature, in which participants are prompted to narrate memories that elicit significant emotional responses (Kensinger, 2009). In contrast, our study prompted participants to reflect on their experiences of feeling a lack of belonging. This more generic prompt may not elicit emotionally vivid memories. In fact, participants' responses may reflect a neutral description of their relevant experiences. Importantly, as the coding constructs carry different meanings across different prompts, the narrative valence does not mean the same thing in different studies (e.g., current work and Lilgendahl et al., 2018). Accordingly, the results are likely to be different. In addition, while we adopted rule-based compound scores to provide more objective sentiment information, a single score may overlook emotional nuances and complexities in the narratives (e.g., contrasted feelings and resolutions). Indeed, some participants in our study indicated they did not have conflicting bicultural experiences, which could have biased the valence scores. In all, the findings for the narrative valence should be interpreted with caution.

Narrative Themes

The second goal of the present study was to identify the themes bicultural individuals use when narrating their experiences. The MEM revealed six themes as follows: cultural belongingness, mixed cultural experience, experience of prejudice, physical characteristics, language experience, and race/ethnic issues.

The first theme we found was cultural belongingness, which encompassed participants' narratives of moving between different cultures. Participants talked about their family's immigration stories and how they try to find belonging as Americans at school while maintaining their heritage cultural traditions at home. Some participants expressed feeling inadequate in their cultural identities or struggling to find belonging. This theme reflects the experiences of many bicultural individuals in the United States, especially considering the country's status as the world's most migratory nation (Budiman, 2020). Participants who mentioned cultural belonging likely feel torn between their heritage and American cultures. This theme aligns with previous findings using open-ended approaches where bicultural individuals shared their uncertainties about their

cultures (Phinney & Devich-Navarro, 1997) and the challenges of being bicultural (Huynh et al., 2018).

The second theme we found was mixed cultural experiences. Bicultural individuals described navigating different cultures in their lives. Participants mentioned how they were raised in their heritage culture while also adopting American ways of socializing. Some participants talked about their identification with a mixed culture, highlighting the emergence of new cultural identities from blended cultures. This theme reflects the complex experiences of bicultural individuals, where new cultural expressions emerge from the blending of their heritage and American cultures. For example, bubble tea culture has emerged as a cultural symbol for Asian Americans, connecting them to heritage cultures and reflecting their unique Asian American identity (X. Zhang, 2021). This theme corresponds to Phinney and Devich-Navarro's (1997) and Huynh et al.'s (2018) findings, in which bicultural individuals discussed their identification with both ethnic and American cultures, feeling of inclusion in American society, understanding of cultural roots, and compatibility of their cultures.

The third theme we found was the experience of prejudice, which captured bicultural individuals' encounters with being labeled as part of a social group that they may or may not identify with. Participants mentioned their experiences of being stereotyped and their feelings of prejudice. As identity is shaped by the construction of past events and the ongoing exploration, bicultural individuals' identity is influenced by how others interact with them and how they feel toward these interactions (McLean & Pasupathi, 2012). The experience of prejudice theme therefore reflects the experiences that shape bicultural individuals' identities.

The fourth theme we found was physical characteristics. Bicultural individuals shared narratives about how their physical features, such as hair color, eye color, and skin tone, influenced their identity. Similar to the experience of prejudice theme, this theme underscores the impact of others' perceptions on bicultural individuals' self-concept. This theme is consistent with Phinney and Devich-Navarro's (1997) findings, in which bicultural individuals based their sense of identity on physical terms (e.g., skin complexion and rough hair).

The fifth theme we found was related to language experience. Participants shared narratives about using their heritage language, both in the United States and in their heritage community. Some participants discussed feeling excluded because of their accents, while others mentioned how language represents their identity. In the United States, more than 20% of residents speak a foreign language (i.e., not English) at home (Zeigler & Camarota, 2019), indicating the prevalence of bilingualism. Given that bilingualism and biculturalism are closely related (nearly 82% of the participants in our study identified as bilingual; Grosjean, 2015; LaFromboise et al., 1993), this theme underscores the significance of language in shaping bicultural identity. This theme also corresponds to Phinney and Devich-Navarro's (1997) findings, indicating that bilingualism enriches bicultural individuals' experiences and influences who they are.

Finally, the sixth theme we found was race/ethnic issues. Different from the physical characteristics theme, this theme specifically focused on experiences related to race and/or ethnicity. Participants narrated their confusion about not fitting into specific racial or ethnic categories and facing rejection from certain communities for not looking "typical" enough. In the United States, race and ethnicity are seen as important identity markers; people categorize others and perceive themselves accordingly (Bonilla-Silva, 2004; Feagin & Ducey, 2018). This theme thus reflects the social dynamics of the context where participants reside.

The language experience and race/ethnic issues themes stood out as two separate factors (i.e., their factor loadings were notably different), suggesting that participants tended to discuss one theme or the other. Generational differences in bicultural identity may have an influence on this distinction. First-generation immigrants often prioritize engagement with the host culture, leading to a link between higher host language proficiency and increased BII (Benet-Martínez et al., 2021). In contrast, second-generation biculturals may prioritize identification with their heritage

culture (Huynh et al., 2018). As they were raised in the dominant mainstream society, second-generation biculturals are also more sensitive to and aware of intergroup concerns such as race/ethnic discriminations (Yoo et al., 2009). Consistent with these findings, we found that U.S.-born bicultural individuals (i.e., second-generation onwards) were less inclined to discuss language experience ($r = -.05$, $p = .29$) but significantly more likely to mention race/ethnic issues ($r = .12$, $p = .02$). However, as the majority of our participants were U.S.-born ($n = 306$, 76.1%), our results may be skewed toward second-generation biculturals. Future research would benefit from exploring a more diverse generational range to better understand the bipolar nature of language experience and race/ethnic issues.

Narrative Valence, Themes, and BII

The third and fourth goals of the present study were to examine the association between narrative valence and BII scales, as well as the themes and BII scales. We focus our discussion on harmony and blendedness instead of considering the total BII for a better interpretation of the findings. The Harmony and Blendedness subscales correlated slightly ($r = .11$, $p = .03$) in this study, indicating that the components are distinguishable, as proposed in previous studies (Benet-Martínez & Haritatos, 2005; Cheng et al., 2014; Huynh et al., 2011, 2018).

Cheng and Lee (2013) found that recalling positive bicultural experiences increases BII, whereas recalling negative bicultural experiences decreases BII. Despite participants in this investigation having relatively positive sentiment when narrating their bicultural experiences ($M = 0.31$, $SD = 0.46$), we found that the narrative valence did not correlate with harmony or blendedness. Our results are explained by Lilgendahl et al.'s (2018) findings, which extended Cheng and Lee's (2013) study. In determining what constitutes a positive bicultural experience, Lilgendahl and colleagues (2018) showed that it is the extent to which narratives ended positively rather than the emotionality of events themselves that is more indicative of BII. The authors found that positive valence on the narrative's endings positively correlated with both harmony and blendedness; however, the overall valence of the text was unrelated to BII. In line with these findings, our results support the idea that the emotionality of bicultural experiences may not be the only determinant factor of BII. Further investigation comparing the narrative's overall valence and ending emotionality would deepen the current understanding.

In regard to the relationship between identity narrative themes and BII scales, our results revealed that the cultural belongingness theme was negatively related to harmony while positively correlated with blendedness. Bicultural individuals who mentioned the cultural belongingness theme more frequently may struggle to fit in with either culture or feel less belonging to the host society. These individuals feel more conflicted with their identities, have a weaker sense of belonging to both cultures, and have more experiences feeling like outsiders. They feel less harmonized between their cultures, thus being lower on harmony, which captures the affective component of bicultural experiences (Benet-Martínez & Haritatos, 2005; Cheng et al., 2014; Huynh et al., 2011, 2018). In contrast, the cultural belongingness theme was positively related to the Blendedness dimension. This is in agreement with previous studies that the blendedness captures the cognitive-behavioral component of identity integration, which involves bicultural competencies such as speaking different languages and practicing cultural traditions (Huynh et al., 2018). With more experiences engaging in both cultures, bicultural individuals become more familiar with dominant and heritage culture practices, contributing to higher blendedness. However, as they become more involved in both worlds, bicultural individuals may feel more conflicted with their identities and feel distant from either culture. They may struggle to find belongingness and narrate the cultural belongingness theme more frequently. Consequently, the more bicultural individuals use the cultural belongingness theme, the less harmony there is and the more blendedness there is.

Similarly, the theme of mixed cultural experience was also negatively correlated with harmony but positively related to blendedness. Past studies found that higher identification with mainstream culture is related to higher harmony (Huynh et al., 2018). It is possible that participants who are more likely to narrate mixed cultural experiences identify less strongly with mainstream American culture and thus mention different cultural experiences more frequently. Accordingly, these individuals may have lower harmony (Huynh et al., 2011). In addition, bicultural individuals who narrate mixed cultural experiences more tend to have more experiences engaging in both cultures. Although they do not perceive their cultural identities as compatible, they may be familiar with both cultures' traditions. Thus, reflecting the cognitive-behavioral component of identity integration, these individuals have higher blendedness.

Regarding the physical characteristics and the language experience themes, our results demonstrated that these themes only correlated with harmony. We found that the more bicultural individuals mention experiences relevant to their physical features, the less they view their identities as compatible. This aligns with Phinney and Devich-Navarro's (1997) findings that alternating biculturals (i.e., those who identify more strongly with their ethnic-cultural group) mention physical traits as part of their bicultural identities. Alternating biculturals perceive their different cultures as less harmonized, which corresponds to our findings that the physical characteristics theme is negatively associated with harmony. Bicultural individuals who perceive their physical features as conflicting with their bicultural experiences may view their different cultural identities as less compatible, so they do not harmonize their different identities.

For the language experience theme, past studies found that bicultural individuals who have more mainstream language barriers tend to have lower harmony (Benet-Martínez & Haritatos, 2005; Miller et al., 2011). Current results show that participants who mention the language experience theme more frequently have lower harmony. As the language experience theme captures participants' practices using their heritage language, those who are less dominant in the mainstream language may have more experiences relating to their heritage language and thus mention this theme more frequently. Accordingly, bicultural individuals who narrate language experiences more frequently tend to have lower harmony. These results also support Phinney and Devich-Navarro's (1997) findings that alternating biculturals (i.e., those who found their cultures conflicting) mention bilingual experiences as part of their bicultural identities.

In sum, our findings suggest that bicultural individuals' identity narratives are positively toned. Bicultural individuals describe their experiences with themes such as cultural belongingness, mixed cultural experience, experience of prejudice, physical characteristics, language experience, and race/ethnic issues. The frequency of mentioning each theme also relates to the extent to which bicultural individuals integrate their identities. In the next section, we discuss the advantages, contributions, limitations, and future directions.

Advantages and Contributions

Identity is about belonging and commonality with others (Weedon, 2004). However, the possible belongings of humans are infinite, especially in bicultural or multicultural populations. For instance, Phinney and Devich-Navarro (1997) stated that multiculturalism is a multidimensional phenomenon; there is not just one way of being bicultural or multicultural. Research in BII has shown that there are pronounced individual differences in how bicultural individuals experience their different identities (e.g., Benet-Martínez & Haritatos, 2005; Cheng et al., 2014). Thus, using scale-level quantitative approach to measure identity is far from enough. This study contributes to the field by using natural language analysis to gain deeper insights into how people narrate their bicultural identity experiences.

More specifically, we adopt the MEM that extracts psychologically meaningful themes from participants' open-ended narratives (Boyd, 2017; Chung & Pennebaker, 2008a, 2008b). The

MEM allows us to define themes in identity narratives that direct individuals' thinking processes. The method avoids imposing constructs, theories, or researcher-predetermined themes as it relies on the most common content words used in participants' narratives. This bottom-up approach allows for themes to emerge from the data, similar to grounded theory, but in an automated manner. Importantly, the MEM enables the analysis of narratives from a larger sample without the need for manual coding, reducing time and resource burdens. Moreover, the MEM is accessible as it requires little coding expertise, and, the exploratory nature of this approach also offers insights that quantitative-based approaches might overlook (Markowitz, 2021). Another unique advantage of the MEM is its suitability to study natural language data in the bicultural or multicultural sphere, where individuals may express their identity in nuanced ways not captured by traditional scale-level approaches. Our study expands the text-analytic toolkit for studying bicultural identity, moving from manual coding by human judges (e.g., Lilgendahl et al., 2018; Syed & Azmitia, 2010) to automated analysis that integrates quantitative and qualitative advantages.

However, one limitation of the MEM is the subjective process of labeling factors, as researchers manually construct labels that encompass most words within a dimension. While the themes extracted from the MEM represent the salient dimensions in the text corpus, they may not encompass all texts. As in the present study, it is possible that participants include a broad range of expressions that could not be classified into the six narrative themes. We addressed this issue by investigating a large sample size—more than 400 bicultural students—to provide adequate power. We also incorporate quantitative measurements (i.e., BII; Huynh et al., 2018) to minimize the uncertainty of using qualitative text-analytic tools. Future work could compare different methods for selecting the number of factors and incorporate different scale-level measurements to further validate the use of natural language analysis.

Limitations and Future Research

While the current study has important real-world implications, there are limitations that stem from the nature of the study methods. First, our sentiment analysis employs the bag-of-words model to classify words as positive, negative, or neutral, thereby determining the narrative valence (Y. Zhang et al., 2010). This may introduce bias when a generic prompt is used. That is, following the prompt in the current study, participants may recall less emotionally intense memories (Kensinger, 2009). This influences how they frame their identity narratives, in which participants may emphasize their experiences rather than focusing on the emotions involved. For instance, our findings suggest that people who narrated cultural belongingness and mixed cultural experience themes had more positive sentiment but scored lower on BII Harmony. This suggests a discrepancy between how individuals perceive their identities and the positive tones they have. To address potential bias, future research could explore different narrative prompts that elicit stronger emotional responses or employ alternative sentiment analysis approaches.

Next, given that the study was administered in English, participants' responses may be biased toward American cultures. As culture and language are closely related, bicultural individuals tend to be bilinguals (Grosjean, 2015; LaFromboise et al., 1993). Indeed, using the MEM, Rodríguez-Arauz et al. (2017) found Mexican American bilinguals think about different self-schemas when primed in different languages. As most of our participants identified as bilingual ($n = 329$, 81.8%), their identity narratives may be influenced by the language they used. Further studies could explore whether language plays a role in bicultural bilinguals' identity narratives by asking participants to respond in different languages.

Moreover, our premise is that identity integration is a generalized concept that describes an individual's experiences across any form of social identity. Thus, we aim to understand the wide range of themes bicultural individuals discuss when they feel they must check multiple boxes. Accordingly, we used an inclusive narrative prompt to encourage exploration of different themes. As participants were unaware that they were recruited based on their bicultural background, nor

did we prompt them to specifically think about their bicultural experiences, we cannot know if participants were thinking about their bicultural identities when narrating their experiences. However, our results showed that even with a generic prompt, bicultural individuals talk about their culture-related experiences, as seen in the cultural belongingness and mixed cultural experience themes. These findings suggest that bicultural experiences are central to bicultural individuals' multiple identities. To further investigate the importance of biculturalism in multiple identity experiences, future research could explore if these results hold when people are specifically asked to reflect on their bicultural experiences.

Finally, we did not find significant links between BII and race/ethnic issues or experiences of prejudice themes. This may seem to contradict previous research that found a link between BII and racial/ethnic prejudice (e.g., Benet-Martínez et al., 2021). Given that a majority of our participants identified as White ($n = 207$, 51.5%), the racial majority in the United States, race/ethnic issues might not be as prominent in their narratives of bicultural experiences, potentially contributing to our findings. Regarding the experience of prejudice, participants recounted diverse narratives related to gender, religion, culture, and race following our generic prompt (see Table 3). Given our framing of bicultural individuals as those identifying with more than one culture, one being American and the other being any other cultural identification, participants may have compared their heritage culture with American culture in their responses to BII. The lack of a significant link between BII and experience of prejudice may be due to the fact that our BII measure may not have fully captured the range of identities of the participants. Because BII has been shown to work with identity dynamics other than culture and ethnicity, our results point the way for more research to improve how we think about biculturalism and look into it in different settings, such as race, class, and gender.

Concluding Remarks

To conclude, bicultural individuals construct unique identity narratives based on their multiple cultural experiences. Our findings, derived from modern natural language analyses, suggest that these narratives are positively toned. They reflect diverse backgrounds, such as cultural belongingness, mixed cultural experience, experience of prejudice, physical characteristics, language experience, and race/ethnic issues. These experiences also relate to the extent to which bicultural individuals perceive their multiple cultural identities as harmonized and blended. In essence, bicultural identity is a dynamic, active, and ongoing process. Bicultural individuals may not belong to one category but can have various identities.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon request.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iD

Merrisa Lin ![ORCID] https://orcid.org/0000-0002-6399-778X

Supplemental Material

Supplemental material for this article is available online.

References

Benet-Martínez, V. (2012). Multiculturalism: Cultural, social, and personality processes. In K. Deaux & M. Snyder (Eds.), *The Oxford handbook of personality and social psychology* (pp. 623–648). Oxford University Press.

Benet-Martínez, V., & Haritatos, J. (2005). Bicultural Identity Integration (BII): Components and psychosocial antecedents. *Journal of Personality*, *73*(4), 1015–1050. https://doi.org/10.1111/j.1467-6494.2005.00337.x

Benet-Martínez, V., Lee, F., & Cheng, C. Y. (2021). Bicultural identity integration: Components, psychosocial antecedents, and outcomes. In *Handbook of advances in culture and psychology* (Vol. 8). https://doi.org/10.1093/oso/9780190079741.003.0006

Benet-Martínez, V., Leu, J., Lee, F., & Morris, M. W. (2002). Negotiating biculturalism: Cultural frame switching in biculturals with oppositional versus compatible cultural identities. *Journal of Cross-Cultural Psychology*, *33*(5), 492–516. https://doi.org/10.1177/0022022102033005005

Berry, J. W. (1990). Psychology of acculturation. In J. J. Berman (Ed.), *Nebraska symposium on motivation, 1989: Cross-cultural perspectives* (pp. 201–234). University of Nebraska Press.

Bonilla-Silva, E. (2004). From bi-racial to tri-racial: Towards a new system of racial stratification in the USA. *Ethnic and Racial Studies*, *27*(6), 931–950. https://doi.org/10.1080/0141987042000268530

Boyd, R. L. (2017). Psychological text analysis in the digital humanities. In S. Hai-Jew (Ed.), *Data analytics in the digital humanities* (pp. 161–189). Springer. https://doi.org/10.1007/978-3-319-54499-1_7

Boyd, R. L. (2018). *MEH: Meaning Extraction Helper* (Version 2.2.03) [Software]. https://www.ryanboyd.io/software/meh

Boyd, R. L., & Pennebaker, J. W. (2016). A way with words: Using language for psychological science in the modern era. In C. V. Dimofte, C. P. Haugtvedt, & R. F. Yalch (Eds.), *Consumer psychology in a social media world* (pp. 222–236). Routledge/Taylor & Francis.

Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings* (Technical Report C-1). Center for Research in Psychophysiology, University of Florida. https://e-lub.net/media/anew.pdf

Budiman, A. (2020). *Key findings about U.S. immigrants*. Pew Research Center. https://www.pewresearch.org/fact-tank/2020/08/20/key-findings-about-u-s-immigrants/

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavior Research*, *1*, 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Cheng, C. Y., & Lee, F. (2013). The malleability of Bicultural Identity Integration (BII). *Journal of Cross-Cultural Psychology*, *44*, 1235–1240. https://doi.org/10.1177/0022022113490071

Cheng, C. Y., Lee, F., Bene t-Martínez, V., & Huynh, Q. L. (2014). Variations in multicultural experience: Influence of bicultural identity integration on socio-cognitive processes and outcomes. In V. Benet-Martínez & Y.-Y. Hong (Eds.), *The Oxford handbook of multicultural identity* (pp. 276–299). Oxford University Press.

Chung, C. K., & Pennebaker, J. W. (2008a). Revealing dimensions of thinking in open-ended self-descriptions: An automated meaning extraction method for natural language. *Journal of Research in Personality*, *42*, 96–132. https://doi.org/10.1016/j.jrp.2007.04.00

Chung, C. K., & Pennebaker, J. W. (2008b). Variations in the spacing of expressive writing sessions. *British Journal of Health Psychology*, *13*, 15–21. https://doi.org/10.1348/135910707X251171

Feagin, J. R., & Ducey, K. (2018). *Racist America: Roots, current realities, and future reparations*. Routledge.

Grosjean, F. (2015). Bicultural bilinguals. *International Journal of Bilingualism*, *19*(5), 572–586. https://doi.org/10.1177/1367006914526297

Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Huynh, Q. L., Nguyen, A. M. D., & Benet-Martínez, V. (2011). Bicultural identity integration. In S. J. Schwartz, K. Luyckx, & V. L. Vignoles (Eds.), *Handbook of identity theory and research* (pp. 827–842). Springer.

Huynh, Q. L., Benet-Martínez, V., & Nguyen, A.-M. D. (2018). Measuring variations in bicultural identity across U.S. ethnic and generational groups: Development and validation of the Bicultural Identity Integration Scale—Version 2 (BIIS-2). *Psychological Assessment*, *30*(12), 1581–1596. https://doi.org/10.1037/pas0000606

Ikizer, E. G., Ramírez-Esparza, N., & Boyd, R. L. (2019). #sendeanlat (#tellyourstory): Text analyses of tweets about sexual assault experiences. *Sexuality Research and Social Policy*, *16*(4), 463–475. https://doi.org/10.1007/s13178-018-0358-5

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion Review*, *1*(2), 99–113. https://doi.org/10.1177/1754073908100432

LaFromboise, T., Coleman, H. L. K., & Gerton, J. (1993). Psychological impact of biculturalism: Evidence and theory. *Psychological Bulletin*, *114*(3), 395–412. https://doi.org/10.1037/0033-2909.114.3.395

Leong, C. H., & Ward, C. (2000). Identity conflict in sojourners. *International Journal of Intercultural Relations*, *24*(6), 763–776. https://doi.org/10.1016/S0147-1767(00)00030-4

Li, J., & Hovy, E. (2017). Reflections on sentiment/opinion analysis. In E. Cambria, D. Das, S. Bandyopadhyay, & A. Feraco (Eds.), *A practical guide to sentiment analysis. Socio-affective computing* (Vol. 5). Springer. https://doi.org/10.1007/978-3-319-55394-8_3

Lilgendahl, J. P., Benet-Martínez, V., Bishop, M., Gilson, K., Festa, L., Levenson, C., & Rosenblum, R. (2018). "So now, I wonder, what am I?" A narrative approach to bicultural identity integration. *Journal of Cross-Cultural Psychology*, *49*(10), 1596–1624. https://doi.org/10.1177/0022022118801555

Liu, B. (2010). Sentiment analysis and subjectivity. *Handbook of Natural Language Processing*, *2*, 627–666.

Markowitz, D. M. (2021). The meaning extraction method: An approach to evaluate content patterns from large-scale language data. *Frontiers in Communication*, *6*, Article 588823. https://doi.org/10.3389/fcomm.2021.588823

McLean, K. C., & Pasupathi, M. (2012). Processes of identity development: Where I am and how I got there. *Identity: An International Journal of Theory and Research*, *12*(1), 8–28. https://doi.org/10.1080/15283488.2011.632363

Meca, A., Eichas, K., Schwartz, S. J., & Davis, R. J. (2019). Biculturalism and bicultural identity development: A relational model of bicultural systems. In P. F. Titzmann & P. Jugert (Eds.), *Youth in superdiverse societies* (pp. 41–57). Routledge.

Miller, M. J., Kim, J., & Benet-Martínez, V. (2011). Validating the Riverside Acculturation Stress Inventory with Asian Americans. *Psychological Assessment*, *23*, 300–310. https://dx.doi.org/10.1037/a0021589

Nguyen, A. M. D., & Benet-Martínez, V. (2007). Biculturalism unpacked: Components, measurement, individual differences, and outcomes. *Social and Personality Psychology Compass*, *1*(1), 101–114. https://doi.org/10.1111/j.1751-9004.2007.00029.x

Pennebaker, J. W., Boyd, R. L., Booth, R. J., Ashokkumar, A., & Francis, M. E. (2022). *Linguistic Inquiry and Word Count: LIWC-22*. Pennebaker Conglomerates. https://www.liwc.app

Phinney, J. S., & Devich-Navarro, M. (1997). Variations in bicultural identification among African American and Mexican American adolescents. *Journal of Research on Adolescence*, *7*(1), 3–32. https://doi.org/10.1207/s15327795jra0701_2

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ramírez-Esparza, N., Chung, C. K., Sierra-Otero, G., & Pennebaker, J. W. (2012). Cross-cultural constructions of self-schemas: Americans and Mexicans. *Journal of Cross-Cultural Psychology*, *43*(2), 233–250. https://doi.org/10.1177/0022022110385231

Rodríguez-Arauz, G., Ramírez-Esparza, N., Pérez-Brena, N., & Boyd, R. L. (2017). Hablo Inglés y Español: Cultural self-schemas as a function of language. *Frontiers in Psychology*, *8*, Article 885. https://doi.org/10.3389/fpsyg.2017.00885

Roehrick, K. (2020). *Vader: Valence Aware Dictionary and sEntiment Reasoner (VADER)* (R Package Version 0.2.1). https://CRAN.R-project.org/package=vader

Stone, P. J., Dunphy, D. C., & Smith, M. S. (1966). *The general inquirer: A computer approach to content analysis*. MIT Press.

Syed, M., & Azmitia, M. (2010). Narrative and ethnic identity exploration: A longitudinal account of emerging adults' ethnicity-related experiences. *Developmental Psychology*, *46*(1), 208–219. https://doi.org/10.1037/a0017825

Warner, G., Antolini, T., Arthur, D., & Yan, J. (2021a). *Boxing back*. https://www.npr.org/2021/02/09/965980941/boxing-back

Warner, G., Antolini, T., Arthur, D., & Yan, J. (2021b). *Our boxes, ourselves*. https://www.npr.org/2021/01/27/961208632/our-boxes-ourselves

Weedon, C. (2004). *Identity and culture: Narratives of difference and belonging*. McGraw-Hill Education.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Yampolsky, M. A., Amiot, C. E., & de la Sablonnière, R. (2013). Multicultural identity integration and well-being: A qualitative exploration of variations in narrative coherence and multicultural identification. *Frontiers in Psychology*, *4*, Article 126. https://doi.org/10.3389/fpsyg.2013.00126

Yoo, H. C., Gee, G. C., & Takeuchi, D. (2009). Discrimination and health among Asian American immigrants: Disentangling racial from language discrimination. *Social Science & Medicine*, *68*(4), 726–732. https://doi.org/10.1016/j.socscimed.2008.11.013

Zeigler, K., & Camarota, S. A. (2019, October 29). 67.3 million in the United States spoke a foreign language at home in 2018. *CIS.org*. https://cis.org/Report/673-Million-United-States-Spoke-Foreign-Language-Home-2018#:~:text=As%20a%20share%20of%20the,of%20all%20foreign%2Dlanguage%20speakers

Zhang, X. (2021, November 15). How Boba became a symbol of "Asian cool." *Sixth Tone*. https://www.sixthtone.com/news/1008960/how-boba-became-a-symbol-of-asian-cool

Zhang, Y., Jin, R., & Zhou, Z. H. (2010). Understanding bag-of-words model: A statistical framework. *International Journal of Machine Learning and Cybernetics*, *1*(1), 43–52. https://doi.org/10.1007/s13042-010-0001-0