Reading Between the Lines: Exploring Narrative Identity through Genre-based Sentiment Analysis of a Corpus of Student-Teachers' Narratives

TEH, DAVID TCHAIKOVSKY BOON ERN
Institut Pendidikan Guru Kampus Bahasa Antarabangsa, Malaysia
david.t@ipgm.edu.my

ABSTRACT

In an educational landscape increasingly dominated by utilitarian-technocratic values and affective austerity, the psychosocial dimension of learner identity is often sidelined. This orientation has marginalised the humanities and social sciences, displacing affect, ethics, and narrative from the centre of human inquiry, thereby prioritising productivity and quantification over reflection and emotional growth, reducing learners to data points rather than meaning-makers. This study seeks to reassert the value of psychosocial and affective meaning-making in education by drawing on Narrative Identity Theory, which views writing as a form of self-narration reflecting the writer's emotional and psychological state. Thirty-six student-teachers enrolled in an undergraduate education programme each produced a narrative text of approximately 1,200 words in a genre of their choice. Corpus linguistic methods were employed (using RStudio packages) to map sentiment trajectories against different genres, using these lexical patterns as proxies for the construction of narrative identity. Results indicate a statistically significant, albeit small (v = .06), association between genre choice and affective sentiment, suggesting that narrative identity is (re)constructed as a continuous process rather than a fixed outcome. The findings, interpreted via Kramsch's and Uryu's "third space" theory and Bakhtin's operationalisation of ventriloquism, show that student-teachers' narratives often challenge normative identity archetypes and pedagogical expectations. The study demonstrates the potential of genre-sentiment analysis as a reflective and insight-generating tool for educators seeking to understand students' affective states and identity negotiation. Future research may extend this work to multimodal storytelling and the development of emotional resilience in teacher education contexts.

Keywords: Genre Analysis; Sentiment Analysis; Learner Corpus; Narrative Identity; NRC Sentiment Lexicon

INTRODUCTION

The current era of rapid technological acceleration has intensified society's collective pivot toward Science, Technology, Engineering, and Mathematics (STEM), inadvertently casting a long shadow over and sidelining the humanities and social sciences (Sarpong, 2025). The systemic privileging of STEM disciplines has fostered a utilitarian epistemology wherein knowledge is increasingly valued for its economic yield rather than for its humanistic or ethical significance, thus relegating the humanities to the periphery of relevance. Resultantly, a subtle but pervasive affective erosion, or emotional austerity, takes place. It points to a phenomenon where systemic devaluation of emotional and psychosocial growth is increasingly common. This austerity produces learners and workers who are encouraged to self-regulate in line with market-friendly dispositions (e.g., resilience, grit) but are disincentivised from exploring emotional complexity. Within educational contexts, this utilitarian turn has reconfigured the learner from a meaning-maker into a data point, where an individual's value is measured by quantifiable performance rather than reflective growth (Joseph, 2025).

Yet, it is precisely these affective and storied dimensions of experience that underpin what it means to learn, teach, and become. It is through narrative that individuals and the collective exercise agency, irony, contradiction, and transformation, so humanity is at risk of losing more than just its grip on reality but also crucially its narrative identity and sense of self. This phenomenon is particularly notable in postcolonial or rapidly modernising societies like Malaysia, where education policies often emphasise STEM to align with global economic needs (Choong, 2022). Within classroom communities of learners, this shift has contributed to the erasure of the learner-as-subject. A phenomenon largely observed in non-Malaysian contexts (Han, 2023), this prevailing educational paradigm casts students as consumers of content or producers of measurable outputs, leaving little room for self-authorship.

More concerningly, the teacher-as-consumer narrative is equally pronounced amongst student-teachers, though this is admittedly far more critical considering that they are not only students of the present but also teachers of the future. While the teacher-as-consumer narrative can be beneficial, for instance, student-teachers adhering to prescribed teaching materials when teaching for the first time (see Chye, 2021), this narrative does come into conflict if and when they prefer greater autonomy or agency (Mark et al., 2020). This conflict is further extended when student-teachers struggle to reconcile their creative, emotional self against their professional, formal teacher persona (Premont, 2022). It is possible, however, that teachers may find comfort through self-expression, for instance, in the form of writing (Parke, 2018).

Therefore, narrative identity theory offers a critical counterpoint to contest the erasure of the learner-as-subject and the fossilisation of the teacher-as-consumer paradigms. It posits that individuals construct their sense of self through internalised and evolving life stories shaped by emotion, culture, and genre (Cuc, 2020; Dunlop et al., 2020). Narrative identity becomes a methodological lens that illuminates how individuals deploy narrative genres to make meaning of their lives. In this study, a learner corpus of open-ended narrative writing was used to explore the narrative identity that represents the ongoing psychosocial negotiation of selfhood amongst student-teachers. This is evidenced through genre selection and emotional lexis in their narrative writing. While learner corpora have been used traditionally for error analysis and monitoring learner progress, recent developments in Natural Language Processing (NLP) and Large Language Models (LLMs) have opened up unprecedented avenues for analysing learner affect (Shaik et al., 2023).

The study of narrative identity through genre-based sentiment analysis thus becomes more than an academic exercise, but rather, an act of epistemic resistance. It reclaims the emotional and narrative dimensions of human experience as legitimate sites of inquiry and challenges the marginalisation of the humanities in an age of datafication. By mapping the affective contours of identity performance, this study recentres the learner as a storied, feeling, and agentive subject. This is also consistent with the fact that students continue to narrativise themselves through essays, projects, journals, and digital posts throughout their schooling. These artefacts are rich with emotional cues and genre markers that reveal the ongoing negotiation of identity, thus allowing researchers and teachers to trace how learners perform, contest, or conform to dominant narratives through emotional and literary registers (Goh, 2019) and challenge the affective austerity of the present age (Jarke & Macgilchrist, 2021). For these reasons, this study leverages sentiment analysis as the primary corpus analysis to explore the narrative identity of student-teachers, as both current students and future teachers, seeing that they are key stakeholders who may exert significant influence on future generations.

THEORETICAL FRAMEWORK

This study pivots on narrative writing as a form of psychosocial and psycho-affective engagement and expression that reflects the individual's personal development. It draws from Narrative Identity Theory as the fundamental theoretical foundation foregrounding identity formation and affective development in creative writing. Contrary to the more reader-centred Reader-Response Theory (see Kunjanman & Aziz, 2021), Narrative Identity Theory proposes that a core identity underlies every narrative, which emerges as a form of self-narration through a process of intrinsic sense-making that is deeply personal (McAdams, 2001). Such a form of self-narration is said to mirror the writer's affective and psychological state, thus providing a locus of positioning oneself against the community and society-at-large (Bamberg, 1997).

From this theoretical lens, writing is seen as a form of self-narration, where meaning is constructed from lived experience, or "life stories", as Ricœur (1991) termed it. However, while McAdams' (2001) operationalisation of self-narration revolved around the notion of explicit autobiography, in this paper, creative narrative writing functions as a "third space" (Kramsch & Uryu, 2020) for identity work. By virtue of opting for specific character archetypes in various narrative genres, students effectively engaged in narrative positioning (Bamberg, 1997), in which they project their internal values and conflicts via fictional characters and events to safely negotiate them. Text production is, therefore, more than just a mechanical production of an encoded message. Instead, it entails a process and snapshot of one's (ongoing) psychosocial and psycho-affective sense-making that never truly ends. Through the writing process, the writer essentially and constantly constructs and deconstructs his/her role and purpose in relation to the community and environment at large, where said narrative identity is but a cross-section of reality captured in a specific space and time.

A more recent reading of Narrative Identity Theory examines the emerging correlations between how the narrator processes notions of sense-making, especially when experiencing life-changing events. Studies have indicated that positioning the self as the central actor in times of hardship and tribulation allows individuals to assume greater personal agency (McAdams & McLean, 2013). From a literary-structural perspective, narrative identity does not unfold in a vacuum but is mediated by available thematic genres, such as *thriller*, *fantasy*, *slice-of-life*, *romance*, to name a few, each carrying emotional expectations and cultural templates. Here, genres are not merely literary classification, but a marker of social action that frames emotional processing. For instance, thrillers provide a medium for individuals to process fear and anxiety, whereas fantasies allow a framework for exploring agency and power. So, the self-reported thematic genres function as the scaffolding for their affective expression.

By identifying genre markers and mapping them to sentiment trajectories, this study examines how identity is told and performed. This framework exposes how certain hybrid genres (e.g., heroic redemption, trauma survival, romantic resolution) shape which identities become legible or valued. Studies have shown that such a positioning has helped to maintain higher levels of mental fortitude and well-being, seeing that individuals are more equipped to process transformational events in life better (McAdams & McLean, 2013).

To empirically capture this affective dimension of self-narration, this study employs sentiment analysis as a proxy for affective stance. This study posits that the recurrence of specific emotional lexis, for instance, a predominance of *fear* or *trust*, is not random but constitutes a narrative signature. By mapping these affective lexical patterns, this study traces how student-teachers position themselves emotionally within their overlapping identities (see Parke, 2018) as

both learners and future educators by treating their lexicon as a trace of the emotional landscape underpinning their narrative identity. This paper, therefore, proposes that the student-teachers' choice of thematic genres, particularly those that are emotionally or psychologically charged, could reveal the affective and psychosocial dimensions of their academic experience when paired with the sentiment expressed in their writing. In querying their narrative identity, this study attempts to provide a snapshot, or more accurately, a cross-sectional screenshot, of how they make sense of their ongoing reality. This approach recognises how creative (narrative) writing offers a projective space in which emotional, cultural, and cognitive preoccupations can surface through genre and sentiment in lieu of the choice of language use.

RESEARCH QUESTION

In lieu of the theoretical framing presented in the preceding section, the research questions and sub-research questions of this study are as follows: *To what extent is there an association between students' genre preferences and their affective sentiment, and what does this association reveal about their narrative identity construction?*

SUB-RESEARCH QUESTIONS

i.   What are the student-teachers' genre preferences in creative writing?
ii.   What are the dominant affective sentiments expressed in the student-teachers' narratives?
iii. Is there any significant association between the student-teachers' genre preferences and their affective sentiment?

METHODOLOGY

SAMPLING

The participants (*n* = 36) were purposively sampled from teacher training institutes in Malaysia, with all of them sharing a common educational background in Bachelor's in Education (B.Ed) specialising in *Teaching of English as a Second Language* (TESL). Arguably, the sample size complicates efforts to generalise findings of this study, but this limitation is offset by the exploratory nature of its objective. Another key consideration is that the participants have high English language proficiency (CEFR B2 and above, with CEFR C1 as minimum exit requirement) as student-teachers undergoing training to be English language teachers in Malaysian primary schools.

Table 1 depicts the participants' demographic information.

TABLE 1. Demographic Variables

Variable	Category	n	%
Ethnicity	Chinese	5	13.9
	Malay	30	83.3
	Missing	1	2.8
Gender	Female	27	75.0
	Male	8	22.2
	Missing	1	2.8
Institute	A	11	30.6
	B	15	41.7
	C	10	27.7
Genre Selection	*Action*	2	5.6
(Self-Selected)	*Coming-of-age*	5	13.9
	Epistolary	2	5.6
	Fantasy	10	27.8
	Romance	2	5.6
	Slice-of-life	5	13.9
	Thriller	10	27.8

A call for participants was put out, where participants can then submit narrative work via Google Form, which also contains rubrics and instructions pertaining to the short story writing. As demonstrated in the demographic data (Table 1), the participants' genre preference was entirely participant-driven. Genre balance was not a key consideration, as ensuring genuine narrative preference in the corpus took precedence. Consequently, the resulting distribution served as a preliminary ecological insight into the participants' reading and viewing habits rather than an artificially balanced dataset. There was no specific prompt provided, as long as submissions are genuinely human and developed based on Freytag's Pyramid Story Plot. Submission was open to all TESL student-teachers studying in Malaysia, which was finally closed (after 40 days) when no new submissions were received for 7 consecutive days. All submissions were anonymised and participants assigned pseudonyms that are only known to the author and individual participants themselves.

CORPUS CONSTRUCTION

The participants were asked to write a creative short story of approximately 1000 to 1200 words in Microsoft Word file format. Crucially, no specific thematic prompt was provided to grant participants full autonomy regarding the subject matter, plot, and character development. This open-ended design was intentional, ensuring that the resulting genre preference and emotional themes truly emerged from their identity negotiation rather than external instructional directives. The established word count also helps ensure sufficient lexical density for robust sentiment analysis. As the task was completed and submitted digitally via Google Form, the participants were free to draft and revise their submission in a self-paced environment distinct from timed, in-class writing. The use of artificial intelligence of any form was prohibited. The short stories were submitted with an explicit declaration that the participants acknowledge and consent to their narrative piece being subjected to analyses and scholastic publications. The short stories were converted and saved in *.txt* format with *utf-8* encoding compliance prior to being imported into RStudio for further analyses. Cleaning and wrangling of raw data files involved resolving line breaks, errors arising from conflicting encryption and encoding, and removal of any images or embedded items. Metadata tagging was conducted in each file in regard to: (i) gender, (ii) ethnicity, (iii) title, and (iv) genre. The author completed the preliminary round of tagging genre types for

each entry before engaging a peer to cross-verify said categorisation. This was critical as genre choice was primarily determined by the participants in writing. To mitigate the ambiguity of idiosyncratic genre labels, narratives were tagged by the author according to standard literary conventions to ensure comparability across the dataset. In this case, participants were not consulted on the final classification, though the genre tagging was member-checked by a faculty member with no direct connection or investment in this paper.

DATA ANALYSIS

Data analysis was conducted via RStudio to run both sentiment and corpus linguistic analyses. The packages are available via the Comprehensive R Archive Network (CRAN) repository. Several analyses were executed:

i.   Sentiment Analysis using NRC Lexicon via "*tidytext*" and "*textdata*" packages
ii.   Keywords in Context (KWIC) query via "*quanteda*" package

The NRC Emotion Lexicon (Mohammad & Turney, 2013) was selected because it provides multidimensional emotional categories rather than simple polarity or scalar valence. This richer mapping of affective meaning allows for a more nuanced analysis of the emotional tone and identity positioning in teacher and learner discourse. Other lexicons, such as AFINN and BING, were considered but were deemed unsuitable as they were limited to positive–negative valence distinctions and lacked the interpretive depth necessary for understanding affective engagement in educational contexts. Moreover, the NRC lexicon works well with the post-hoc Chi-Square Test of Independence when bootstrapping the association between genre and sentiments.

The use of NRC Emotional Lexicon thus provides the study with a robust framework to look at the sentiments underlying the narrative writing, i.e. investigating complex emotional arcs involving female characters in Malay folktales (see Ramlan et al., 2025). However, a major limitation of the lexicon is its inability to detect polysemy, especially when devoid of context. To mitigate this, qualitative spot-checks using Keywords-in-Context (KWIC) were conducted to verify that the sentiment tagging matched the contextual usage via human tagging and verification. From the corpus, the 15 most frequently occurring non-stop words were identified and ranked in order to focus the analysis on the most salient lexical items prior to coding them according to the sentiments expressed in their contexts. For each of these words, the KWIC analysis was conducted using a window of eight tokens on either side, producing individual tables showing each occurrence along with its surrounding text. These KWIC tables were exported as *.csv* files to allow a manual but structured examination. Each occurrence was then interpretively annotated for sentiment, with an additional column "contextual meaning" added to aid tagging. For instance, the token "*time*" was mapped against the sentiments *anticipation* and *disgust* based on the sentiments evoked from their contextual use (Table 2):

TABLE 2. Kwic "Time" Mapped Against Its Corresponding Sentiments

Sentiment	KWIC for "*time*"
anticipation	*Panic set in. We searched everywhere. **Time** was ticking. the gate would close in...*
disgust	*...it's all anyone talks about. **Time** for your meds. **Time** for tests. And how much **time** do you have left?*

After all the annotations were completed and reviewed via member-checking, the KWIC tables were re-imported into RStudio and combined into a single dataset, thus retaining both the lexical occurrences and their human-tagged sentiments. This combined data set was then grouped by keywords and sentiments to calculate the frequency of each word within passages corresponding to each sentiment category. Finally, for each sentiment, the top three most frequent keywords were identified, with each keyword presented alongside its occurrence usage (see

Table 3). This workflow ensured that sentiment assignment was sensitive to contextual usage, preserved transparency through CSV-based KWIC records, and combined quantitative frequency analysis with qualitative, context-aware manual annotation, despite being time-consuming and labour-intensive. This manual validation helped confirm that the lexicon largely aligned with the participants' intended meaning in their creative writing.

FINDINGS

RQ1: WHAT ARE THE STUDENT-TEACHERS'GENRE PREFERENCES IN CREATIVE WRITING?

Figure 1 depicts the participants' genre preference in their creative writing. The participants largely preferred writing *fantasy*, n = 10, (including *magic*, *fairy tales*) and *thriller*, n = 10, (including *horror*, *supernatural*), the most. *Coming-of-age* and *slice-of-life* are the next most frequent at five each, with *action*, *epistolary*, and *romance* tallying at two each. This finding suggests clear clustering in terms of the participants' genre selection, albeit from a relatively small sample.
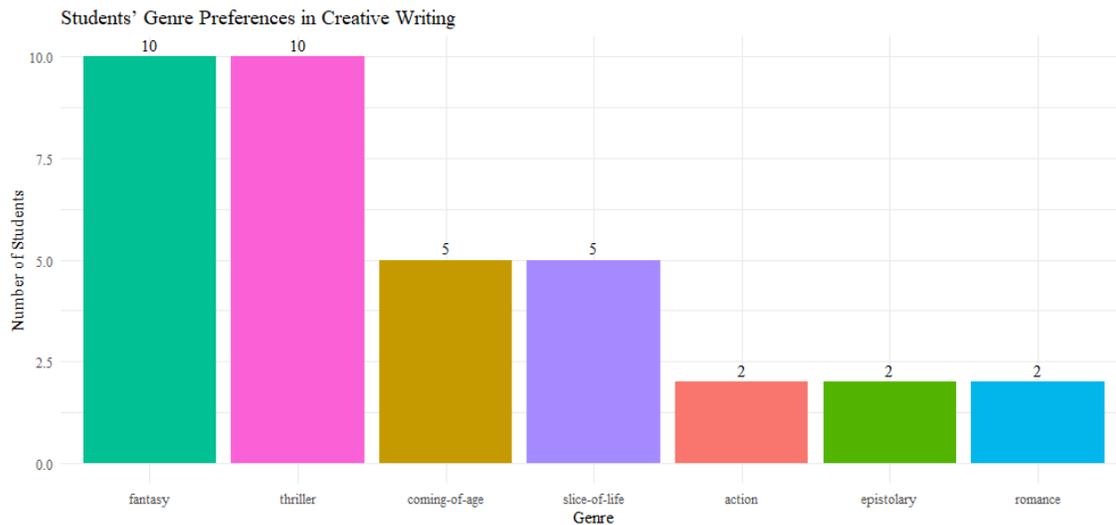

FIGURE 1. Participants' Genre Preferences In Creative Writing

RQ2: WHAT ARE THE DOMINANT AFFECTIVE SENTIMENTS EXPRESSED IN THE STUDENT-TEACHERS' NARRATIVES?

Figure 2 portrays the sentiment distribution in the corpus. The dominant sentiments emerging from the corpus were the valence emotions, namely positive and negative. Negative sentiment (n = 558) slightly edged ahead of positive sentiment (n = 510), suggesting an overall near equilibrium between positively and negatively connoted diction across all samples. This lexical symmetry implies that participants tend to balance affective polarities, perhaps reflecting the nuanced emotional awareness characteristic of young adult writers who negotiate between optimism and realism. Among the eight discrete emotions, *fear* (n = 319), *trust* (n = 270), and *sadness* (n = 262) were the most prevalent, followed by *anger*, *anticipation*, *disgust*, *surprise*, and *joy* in descending order. The relative prominence of *fear* and *sadness* is expected, given the participants' genre preferences, particularly the *thriller* and *fantasy* narratives that often invoke themes of tension, conflict, and existential uncertainty. Conversely, the moderate presence of *trust* suggests that, despite the darker tonalities of their narratives, participants still embedded relational and moral dimensions within their storytelling.
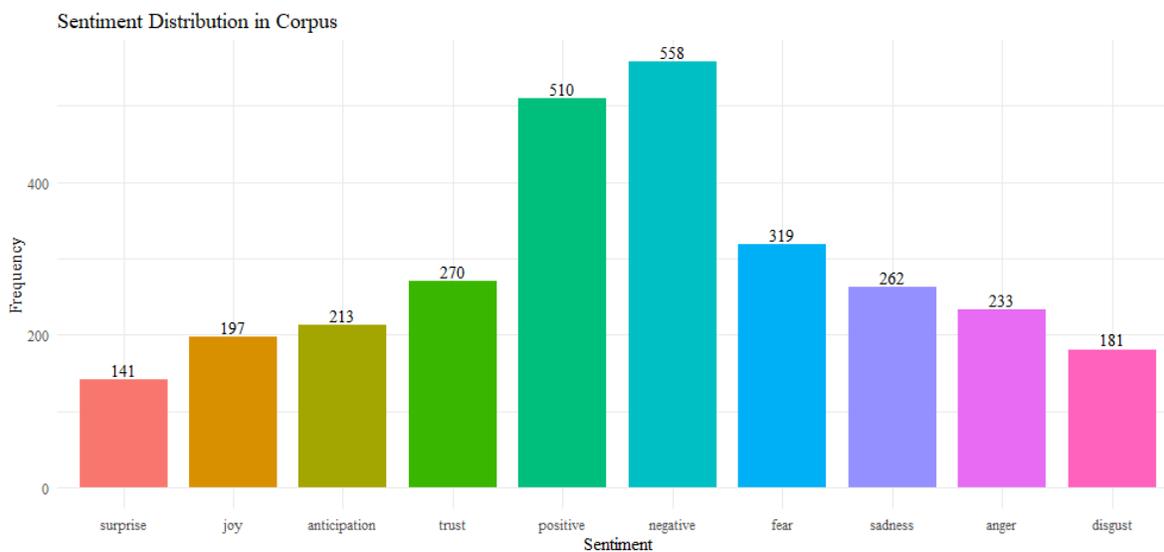

FIGURE 2. Bar Chart Of Sentiment Distribution In Corpus

Table 3 depicts the top three most frequently observed tokens mapped against their corresponding sentiments using KWIC analysis. A closer lexical inspection of each sentiment's top keywords reinforces these patterns. Words associated with negative emotions, such as "*cold*," "*broken*," and "*smile*", indicate preoccupations with discomfort, interpersonal strain, and mitigated positivity, themes commonly aligned with adversity-driven or challenging narrative moments. Positive sentiments, by contrast, were characterised by lexical items like "*finally*", "*smile*", and *"tree*", denoting resolution, affirmation, and growth. These are often key elements commonly found in self-actualising or reconciliatory storylines. Similarly, *fear* was frequently evoked through words like "*fear*", "*dark*", and "*cold*", reflecting anxiety surrounding uncertainty or potential threat rather than overt panic. The *trust* cluster, containing words such as "*mother*", "*found*", and "*quiet*", underscores the importance of relational anchors and stability as moral or

emotional touchstones, even within tension-filled narratives. On the other hand, *sadness* shared lexical overlaps with *fear* and negative sentiments, centring around words such as "*time*", "*words*", and "*quiet*", which evoke reflection, absence, or subdued affect. Less dominant but still notable were *anticipation* and *joy*, featuring tokens like "*time*", "*thought*", and "*finally*" which collectively suggest forward movement, expectation, and small but meaningful moments of satisfaction, albeit in smaller proportion, compared to darker affective tones.

TABLE 3. Top Three Most Frequent Keywords Corresponding to Each Sentiment According to Contextual Use

No.	Sentiment	Top 1	Top 2	Top 3
1	anger	words (6)	time (2)	cold (1)
2	anticipation	time (42)	thought (22)	finally (7)
3	disgust	time (4)	cold (2)	finally (1)
4	fear	fear (19)	dark (16)	cold (10)
5	joy	found (13)	time (12)	smile (11)
6	negative	cold (15)	broken (12)	smile (10)
7	positive	finally (15)	smile (10)	tree (9)
8	sadness	time (16)	words (11)	quiet (7)
9	surprise	suddenly (20)	found (8)	thought (3)
10	trust	mother (14)	found (12)	quiet (10)

Note. The figures in parentheses represent the frequency of each word as contextualised by the corresponding sentiments

RQ3 IS THERE ANY SIGNIFICANT ASSOCIATION BETWEEN THE STUDENT-TEACHERS' GENRE PREFERENCES AND THEIR AFFECTIVE SENTIMENT?

A heatmap was generated to map the intersections between sentiments and genre types in the corpus.

Figure 3 depicts the heatmap showing greater coverage on both valence sentiments, followed by small pockets of coverage on independent scalar sentiments in different genres. To ensure comparability across narratives of varying lengths, sentiment frequencies were normalised to per 1,000 words.
Figure 4 visualises this relationship, highlighting sentiment means in the top quartile to identify dominant emotional signatures.
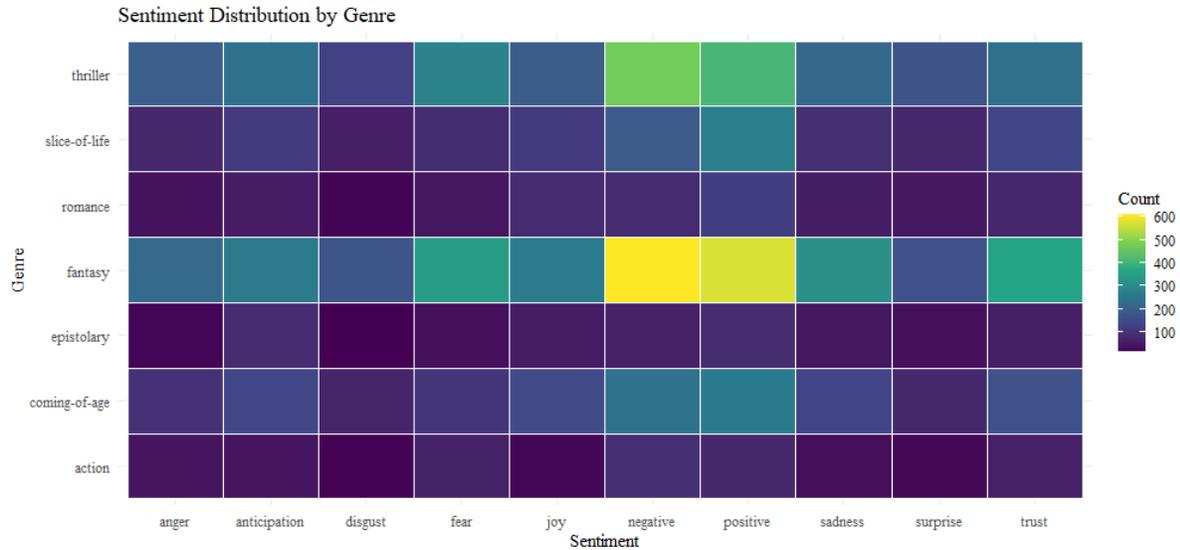
FIGURE 3. Sentiment Distribution Heatmap By Genre

FIGURE 4. Targeted Visualisation Of Sentiment Distribution By Genre

The analysis reveals distinct affective profiles: valence sentiments aside, a large number of genres, such as "*romance*", "*fantasy*", "e*pistolary*", exhibit monotonous emotional range. Interestingly, valence sentiments dominate in *thrillers*–one would be forgiven to think that *fear* or *anticipation* would be a dominant sentiment of that genre. Further analysis revealed several nuanced patterns. Highlighting the top quartile of normalised sentiment frequencies across genres enabled data reduction and clarified the distribution of salient affective tendencies. The visualisations show that *romance* and *fantasy* genres record comparatively higher mean values for *joy*, *positive*, and *trust* sentiments. In contrast, *action* and *thriller* genres demonstrate greater intensities in *fear* and *negative* sentiments. The *epistolary* genre records the highest *anticipation* mean among all genres, while *slice-of-life* and *coming-of-age* genres exhibit moderate distributions across most sentiment categories. These patterns suggest differentiated sentiment

profiles across genres, with notable concentration of positive emotions in affective genres and negative emotions in suspense-oriented genres.

Further Chi-square test of independence revealed a significant association between genre and sentiment ($\chi^2 = 212.17$, $df = 54$, $p < .01$, $v = .06$). However, the effect size was small (*Cramer's V* $= .06$), suggesting that while there is a detectable relationship between genre choice and emotional expression, it is likely subtle and thus non-deterministic. So, in order to deconstruct these tendencies and identify specific drivers of said association, a post-hoc analysis of standardised residuals was conducted. This approach isolates specific genre-sentiment pairings that significantly deviated from expected frequencies, where residual values of ±2.0 or more are flagged. The association of significant deviations acts as a unique affective signature observed in the corpus, thereby highlighting where the strongest associations lie despite the overall modest effect size. Table 4 outlines the overrepresented and underrepresented sentiments in each genre, henceforth presenting a more holistic view of how the participants' narrative writing was underlaid by their genre of choice.

TABLE 4. Post-Hoc Analysis Of Chi-Square Test Of Independence Between Genres And Sentiments

	Genre	Overrepresented (Strong Positive Association)	Underrepresented (Strong Negative Association)
1	*Action*	*Trust* (+5.85)	*Disgust* (−2.46)
2	*Romance*	*Sadness* (+4.43) *Anticipation* (+2.11)	*Surprise* (−2.65) *Anger* (−2.12)
3	*Fantasy*	*Disgust* (+4.00) *Surprise* (+2.16)	*Joy* (−2.56) *Trust* (−2.54) *Sadness* (-2.80)
4	*Epistolary*	*Anticipation* (+3.15) *Sadness* (+3.15) *Trust* (+2.52) *Joy* (+2.18)	*Disgust* (−3.33) *Surprise* (−3.27) *Positive* (−3.15)
5	*Thriller*	*Disgust* (+2.24)	*Trust* (−3.19)
6	*Coming-of-age*	—	*Anger* (−3.16) *Surprise* (−2.08)
7	*Slice-of-life*	—	*Trust* (−2.63)

Note. Standardised Residuals indicate a statistically significant association, where:
$|2.00| \rightarrow p \approx .05$
$|2.60| \rightarrow p \approx .01$
$|3.00| \rightarrow p \approx .003$,
$|3.3+| \rightarrow p < .001$.

DISCUSSION

The findings of this study have highlighted several nuanced observations emerging from the sentiment-genre pairings by leveraging two key interpretive lenses. The first is Kramsch and Uryu's (2020) concept of "Third Space", where the act of writing functions as a symbolic borderland (Kramsch & Uryu, 2020). In this context, the narrative text mediates between the participants' "first space"–their immediate reality as student-teachers–and the "second space" of their emerging professional identities as authoritative and dutiful future teachers. Within this liminal zone, the participants engage in what Bakhtin termed as ventriloquism (Yan et al., 2023), which is the second key interpretive lens of this paper. Far from mere mimicry, ventriloquism is

understood as a strategic act of appropriating the voice of the Other to negotiate a multi-voiced self (Yan et al., 2023). Combined with Narrative Identity Theory, this study posits that the participants are ventriloquising various narrative and character archetypes throughout the corpus. The following sub-sections shall outline the key findings that emerged from this study.

FINDINGS 1: DISRUPTION AND DESTABILISATION OF NORMATIVE CHARACTER AND NARRATIVE ARCHETYPES

Firstly, the participants have demonstrated a preference for unconventional character and narrative archetypes in their narrative writing, resulting in a gravitation toward hybridised narrative tropes such as reluctant and anti-heroes and dark fantasy. This is notable from the standardised residual analysis, were distinct emotional tendencies manifested across genres, suggesting both expected and unexpected affective alignments. Some genres, such as *coming-of-age*, *epistolary*, *romance*, and *slice-of-life*, yielded results consistent with their conventional emotional profiles. In these categories, positive sentiments such as *anticipation*, *joy*, and *trust* were present, overrepresented even, while negative emotions such as *anger*, *disgust*, and *surprise* were unsurprisingly underrepresented. These affective distributions may reflect the participants' own demographic and psychosocial profile: being in their early twenties and recently transitioning out of tertiary education, they are likely to hold optimistic worldviews, tempered by self-discovery and emerging adulthood. The emotional equilibrium observed across these genres also hints at a degree of affective maturity, where participants express hope and affection without resorting to extreme emotional polarities.

In contrast, *thriller*, *action*, and *fantasy genres* exhibited more complex and counterintuitive emotional signatures. *Disgust* (r_{ij}=+2.24) emerged as a dominant emotion in *thrillers*, accompanied by a significant underrepresentation of *trust* (r_{ij}=-3.19). This imbalance suggests that participants tend to associate *thrillers* with moral ambiguity, betrayal, or revulsion rather than mere *fear*, which is a reflection of contemporary narrative sensibilities that privilege psychological tension over straightforward terror. This pattern may also mirror the participants' own social perception, where *trust* becomes precarious in high-pressure environments such as early professional life as an educator.

The *action* genre produced an intriguing finding: *trust* (r_{ij}=+5.85) was markedly overrepresented, whereas *disgust* (r_{ij}=-2.46) was underrepresented. Contrary to expectations of chaos or aggression, the prevalence of *trust* implies a narrative framing built around camaraderie, loyalty, and moral integrity. This specific sentiment profile parallels the "Reluctant Hero" archetype, a protagonist driven by duty and sacrifice rather than glory (often typified in pop culture by figures such as *Frodo Baggins*, *Katniss Everdeen*, or *Paul Atreides*). While the participants did not explicitly cite these characters, their sentiment usage mirrors the moral architecture of these narratives. This suggests that the student-teachers may be projecting an "aspirational ethic", internalising virtues of responsibility and service over individual triumph.

Similarly, *fantasy* narratives were dominated by unexpected sentiments, such as *disgust* (r_{ij}=+4.00) and *surprise* (r_{ij}=+2.16), with *joy* (r_{ij}=-2.56), *trust* (r_{ij}=-2.54), and *sadness* (r_{ij}=-2.80) notably underrepresented. The prevalence of *disgust* and the scarcity of sentiments such as joy and *trust* represent an unexpected divergence from the idealism of traditional fairy tales. Instead, this emotional profile aligns better with the contemporary trend of retelling darker fantasies that foreground moral complexity and existential struggle. Such stories—reminiscent of the *Brothers Grimm*, *The Lord of the Rings*, or *The Witcher*—reject utopian closure in favour of introspection

and endurance. By gravitating toward these affective structures, the participants appear to be rejecting simplified "happily-ever-after" in favour of narratives of endurance and persistence. This reflects a "generational orientation" toward resilience, where hope is portrayed not as given, but as a hard-won outcome in an increasingly uncertain world. Again, assumptions point to the complex process of ventriloquising the "other" within the liminal "third space".

FINDINGS 2: PREVALENCE OF COUNTER-INTUITIVE EMOTION-GENRE PAIRINGS

Secondly, the findings of this study suggest that while youthful optimism pervades certain genres, the participants' engagement with darker emotional landscapes (especially in *thriller* and *fantasy* narratives) reflects a nuanced affective literacy. The coexistence of *trust* and *disgust*, of *anticipation* and *sadness*, illustrates a generational storytelling voice that balances idealism with disillusionment—an emotional duality characteristic of early adulthood and the formative years of professional identity development. The prevalence of these non-conventional affective archetypes parallels the participants' likely exposure to intertextual pop culture. While direct media consumption data was not collected, the narrative patterns observed in this study, especially the "reluctant hero" and "dark fantasy" tropes, strongly resemble the complex emotional arcs found in modern transmedia storytelling (Busselle & Bilandžić, 2009). If so, the participants are not merely consuming these media narratives but are actively re-voicing them to make sense of their own identities. Also, said exposure would mirror and influence how they perceive reality and project their narrative identity as they interact and socialise with other actors within a shared reality (or community). Past studies have leveraged narrative engagement, in the form of emotional arcs analysis and plot progressions, to track sentiments and reactions to different genres (Schmidt et al., 2023). Seeing that sentiments change and individuals mature over time due to accrued experience, narrative writing has become an observatory of how identity coherence and emotional maturity take place in individuals (McAdams, 2001; McLean et al., 2021).

In the context of this study, the reluctant, anti-hero trope adopted by most participants points to two potential positionings, that: (i) *existing is circumstantial*, but (ii) *being is a choice*. These serve as a reminder that the execution of one's duties as a student-teacher, as in the case of the participants, can be challenging and even discouraging at times, but embracing and committing to the teaching vocation is ultimately a matter of choice. By permitting these character archetypes to emerge via empowerment of the participants' agency and narrative identity, we challenge and demystify dominant discourses surrounding the need for learners, which in this case, student-teachers, to embrace the perceptions of infallibility as future teachers. Therefore, underlying the prevalence of non-conventional character archetypes and hybrid genres is the understanding that the participants, as undergraduate tertiary students, are often situated in a highly tense and pressurised environment. Their circumstance, to a certain extent, is associated with and reflected in the rejection of a utopian worldview. Said rejection also includes processing negative sentiments arising from undesired processes or outcomes. Genre studies have demonstrated how individuals framed traumatic life experience via dystopian/utopian genres (Breheny & Stephens, 2023), which explains why older individuals tend to enjoy darker themes and mood as they accrue more experience in processing negative sentiments and emotions. Moreover, Lee (2021) observed how genre traditions may codify collective affective processing, a point that is largely observed and reproduced in this study. The participants of this study, therefore, likely project their narrative identity through acts of creative storytelling (i.e. genre and sentiment selection), hence facilitating emotional exploration and trauma negotiation.

FINDINGS 3: THE INFLUENCES OF TRANSMEDIA INFLUENCES, INTERTEXTUALITY, AND CONSUMPTION

Thirdly, also emerging from the study is a telling description of the participants' preference and media consumption, resulting in an affective processing that strongly resonates with the notions of "narrative authorship", writer voice and identity (De Mieroop, 2020). For instance, the participants' genre preference in their writing suggests that they probably received significant exposure towards *fantasy* and *thriller* through their interaction with mass media, social media, and (computer and console) gaming. Likewise, the emergence of story-driven and plot-heavy interactive games may explain why a number of stories were written in the *coming-of-age* and *slice-of-life* genres. Considering these assumptions, the participants, through narrative writing, may have engaged in the (de)construction of their (projected) narrative identity, manifesting in their interplay and decision-making associated with their choice of genres and sentiments underlying their narrative writing. This resonates well with the theoretical framework of this paper, which proposed that the participants, as student-teachers, may manifest the results of the processing of their current worldviews co-constructed through their narrative identity. Their preferences for anti-heroes and dark fantasies, over other genre or character hybrids, suggest a rejection of utopian worldviews in an increasingly brittle, anxious, non-linear, and incomprehensible (BANI) world.

Centring their narration in non-conforming narrative hybrids thus resonates closely with conflicts arising from cultural and identity negotiation (Truman, 2016). Perhaps mirroring their choice of genres, the participants demonstrated overlapping emotional states, also known as "third-space", in their written production (Kramsch & Uryu, 2020). This hybridity should not discount the fact that the process of attaining an equilibrium is dynamic and volatile, as previous studies in second language writing have reported severe academic detachment between the second language learners' emerging narrative identity and their narrative product (Pavlenko, 2005). Rahimivand and Kuhi (2014) extended this investigation further amongst tertiary learners, where they wrestled between embracing greater autonomy and agency in their academic work. These studies highlight the potency in analysing the resonance between the participants' choice of genre and the sentiments emerging in their academic work.

Finally, the very nature of the writing task itself may have permitted and successfully facilitated genuine and authentic storytelling from the participants, despite the well-reported alienation and suppression of personal voice in genre-based pedagogy (see Hyland, 2008; Swales, 1990). This stance has bred the confidence that the emergence of non-conventional character archetypes observed in the participants' stories is a partial representation, or at the very least, an extension of their existing psychosocial state, as the final product is something that the authors could take ownership of (Pourdana & Tavassoli, 2022). As the writing task was also framed as a piece of creative writing, some participants sought to challenge existing rigidity in narrative norms, which is evident from the non-conventional character archetypes and hybridity of their choice of genre. This positioning augurs well with Mahoney's (2014) observation of second language graduate students expressing frustration but also demonstrating reinvention of their narrative selves through writing.

As student-teachers, the findings provided an insight into their readiness to step into an increasingly challenging school environment. In an era where their ability as (novice) teachers to process desired and less desired emotions draws greater attention and scrutiny, the value of a genre-based sentiment analysis cannot be overstated. Through writing, student-teachers are provided with more than just a channel to process their emotions–narrative writing thus becomes a platform on which they formulate their (future or ongoing) teaching identities and malleable pedagogical practices.

CONCLUSION

This study has asserted, through a combined genre-sentiment analysis of learner corpus, that the student-teachers' narrative identity was (de)constructed dynamically and often non-conventionally. Examination of their narrative identity was made possible by a dual-focal investigation and triangulation of their genre preference with affective sentiments throughout their narrative writing. The findings have demonstrated, first and foremost, that learners' narrative identity may not always fit the existing archetypes. In contesting and challenging existing social and narrative norms, they actively ventriloquise complex archetypes, such as the "reluctant hero" or "dark fantasy" trope, to negotiate the "third space" between their personal and professional identities (Kramsch & Uryu, 2020).

Methodologically, this study contributes to the field of learner corpus research by demonstrating that sentiment analysis–typically reserved for large-scale opinion mining–can be effectively adapted for small-scale, classroom inquiries. By pairing the NRC Emotion Lexicon with genre analysis, this study offers a novel framework for "reading" the psychosocial and psycho-affective undercurrents of learner texts that might be invisible to or overlooked by traditional rhetorical analysis.

Pedagogically, the findings underscore the value of creative writing as a site of identity construction and formation. The emergence of hybrid and dark genres suggests that learners, when granted agency, process the emotional weight of their future projected self. More specifically, teacher education programmes should consider moving beyond purely functional academic literacy to include expressive "third spaces", where student-teachers can engage and process "taboo" emotions that accompany professional be(com)ing in a safe space.

However, a critical caveat remains. This study is not Freudian by any means, so it is never the intention of this paper or any of the methodology described to diagnose or provide a prognosis to any of the participants' psychosocial or psycho-affective state. The presence of "darker" themes or negative sentiments thus reflects a narrative strategy rather than a clinical psychological state. Likewise, many of the assumptions proposed in this study could not be triangulated without incurring in-person, potentially intrusive, investigation that this study is not equipped nor consented to do. As such, these insights should be used to foster dialogue and empathy, never to pathologise the participants.

Nonetheless, future studies employing a similar epistemological and methodological stance can examine how non-traditional academic genres legitimise and empower taboo or undesired emotions. Likewise, we can also ponder if genre mastery correlates with emotional resilience in marginalised learners, as their narrative identity is likely to manifest in their storytelling. Also, future studies can explore if this methodological approach can be extended to investigate multimodal and increasingly trans-modal narratives. Ultimately, by (re) centring the affective self, this study argues for a more humane, affectively literate approach to better (mutual and self)

understanding, acceptance, empowerment of the individual and of the masses, and action to effect change in both thought and practice.

ACKNOWLEDGEMENT

DECLARATION

Parts of the findings reported in this paper were previously presented at the Malaysian Association of Applied Linguistics International Conference (MAALIC) 2025.

REFERENCES

Bamberg, M. (1997). Positioning between structure and performance. *Journal of Narrative and Life History*, *7*(1–4), 335–342.

Breheny, M., & Stephens, C. (2023). Experiences of COVID-19 lockdown among older people in Aotearoa: Idyllic or dystopian? *Ageing and Society*. https://doi.org/10.1017/s0144686x23000545

Busselle, R. W., & Bilandzić, H. (2009). Measuring narrative engagement. *Media Psychology*. https://doi.org/10.1080/15213260903287259

Choong, C. (2022). Race-based affirmative action in Malaysia: Misrecognised subjectivities, enduring inequalities. *Asian Journal of Social Science*, *50*(2), 79–86.

Chye, S. Y. L. (2021). *Becoming a teacher: Reflection, inquiry and identity*. Office of Teacher Education, National Institute of Education, Nanyang …. https://repository.nie.edu.sg/server/api/core/bitstreams/eb388280-7b75-4ba9-820d-3f0306987248/content

Cuc, M. C. (2020). *Narrative identity – A fundamental concept In the development of the self*. 685–693. https://doi.org/10.15405/epsbs.2020.06.71

De Mieroop, D. V. (2020). The narrative dimensions model and an exploration of various narrative genres. *Narrative Inquiry*. https://doi.org/10.1075/ni.19069.van

Dunlop, W. L., Harake, N., Wilkinson, D., & Graham, L. E. (2020). On rhetoric and ratings: Assessing narrative identity via conceptual coding and self-ratings. *Journal of Research in Personality*, *85*, 103922. https://doi.org/10.1016/j.jrp.2020.103922

Goh, P. (2019). Implementing narrative-pedagogical approaches in a teacher education classroom. *The Qualitative Report*. https://doi.org/10.46743/2160-3715/2019.3869

Han, X. (2023). Subjectivity as the site of struggle: Students' perspectives toward sino-foreign cooperation universities in the era of discursive conflicts. *Higher Education*, *85*(2), 399–413. https://doi.org/10.1007/s10734-022-00840-w

Hyland, K. (2008). Writing theories and writing pedagogies. *Journal on English Language Teaching*, *4*(2), 1–20. https://doi.org/10.25170/ijelt.v4i2.145

Jarke, J., & Macgilchrist, F. (2021). Dashboard stories: How narratives told by predictive analytics reconfigure roles, risk and sociality in education. *Big Data & Society*, *8*(1), 20539517211025561. https://doi.org/10.1177/20539517211025561

Joseph, J. (2025). The algorithmic self: How AI is reshaping human identity, introspection, and agency. *Frontiers in Psychology*, *16*, 1645795. https://doi.org/10.3389/fpsyg.2025.1645795

Kramsch, C., & Uryu, M. (2020). Intercultural contact, hybridity, and third space. In J. Jackson (Ed.), *The Routledge Handbook of Language and Intercultural Communication* (p. 15).

Kunjanman, S., & Aziz, A. (2021). Reader-Response Theory: A systematic literature review. *Malaysian Journal of Social Sciences and Humanities*. https://doi.org/10.47405/mjssh.v6i4.747

Lee, H. (2021). A study on the narrative analysis of Gwangjang through Narrative Identity Theory and Tragedy Theory. *Journal of Humanities and Social Sciences*, *12*(2), 1313–1328. https://doi.org/10.22143/hss21.12.2.93

Mahoney, S. Kathleen. (2014). *A critical review of genre based instruction in a writing and research course for second language graduate students*. 1–107.

Mark, S., Id-Deen, L., & Thomas, S. (2020). Getting to the root of the matter: Pre-service teachers' experiences and positionalities with learning to teach in culturally diverse contexts. *Cultural Studies of Science Education*, *15*(2), 453–483. https://doi.org/10.1007/s11422-019-09956-5

McAdams, D. P. (2001). The psychology of life stories. *Review of General Psychology*, *5*(2), 100–122. https://doi.org/10.1037/1089-2680.5.2.100

McAdams, D. P., & McLean, K. C. (2013). *Narrative identity*. https://doi.org/10.1177/0963721413475622

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, *29*(3), 436–465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Parke, E. (2018). Writing to heal: Viewing teacher identity through the lens of autoethnography. *The Qualitative Report*. https://doi.org/10.46743/2160-3715/2018.3548

Pavlenko, A. (2005). *Emotions and Multilingualism*. Cambridge University Press.

Pourdana, N., & Tavassoli, K. (2022). Differential impacts of e-portfolio assessment on language learners' engagement modes and genre-based writing improvement. *Language Testing in Asia*. https://doi.org/10.1186/s40468-022-00156-7

Premont, D. (2022). 'I couldn't sleep': Tension in preservice teacher writing instruction. *Journal of Language and Literacy Education*, *18*(1), n1. https://files.eric.ed.gov/fulltext/EJ1357978.pdf

Rahimivand, M., & Kuhi, D. (2014). An exploration of discoursal construction of identity in academic writing. *Procedia - Social and Behavioral Sciences*, *98*, 1492–1501. https://doi.org/10.1016/j.sbspro.2014.03.570

Ramlan, W. N. M., Idrus, M. M., & Hassan, N. I. (2025). "No longer languished in the periphery of existence:" Female bodies, computational-sentiment analysis, and Malay folktales. *3L The Southeast Asian Journal of English Language Studies*, *31*(3), 375–395. https://doi.org/10.17576/3L-2025-3103-23

Ricœur, P. (1991). Narrative identity. *Philosophy Today*, *35*(1), 73–81. https://doi.org/10.5840/philtoday199135136

Sarpong, J. (2025). Funding cut for the humanities in neoliberal higher education systems: A case of New Zealand's Marsden Fund. *New Zealand Journal of Educational Studies*, *60*(1), 265–278. https://doi.org/10.1007/s40841-025-00388-0

Schmidt, M.-L. R. C., Winkler, J. R., Appel, M., & Richter, T. (2023). Tracking emotional shifts during story reception: The relationship between narrative structure and affective responses. *Scientific Study of Literature*. https://doi.org/10.61645/ssol.177

Shaik, T., Tao, X., Dann, C., Xie, H., Li, Y., & Galligan, L. (2023). *Sentiment analysis and opinion mining on educational data: A survey*. https://doi.org/10.48550/ARXIV.2302.04359

Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Syed, M., Kate C. McLean, McLean, K. C., David Dunlap, Dunlap, D., Sarah C. Jennings, Jennings, S. C., Nicole Litvitskiy, Nicole S Litvitskiy, Nicole S. Litvitskiy, Jennifer Pals Lilgendahl, & Lilgendahl, J. P. (2021). Stability and change in autobiographical reasoning: A 4-year longitudinal study of narrative identity development. *Journal of Personality*. https://doi.org/10.1111/jopy.12669

Truman, S. E. (2016). Becoming more than it never (actually) was: Expressive writing as research-creation. *Journal of Curriculum and Pedagogy*, *13*(2), 136–143. https://doi.org/10.1080/15505170.2016.1150226

Yan, D., Bright, D., Prosser, H., & Poole, A. (2023). Ventriloquism as method: Writing differently and thinking philosophically. *Cultural Studies ↔ Critical Methodologies*, *23*(3), 262–273. https://doi.org/10.1177/15327086221144789