



LALICO Journal of Languages, Literature, and Communication

Print ISSN: 2651-6462

www.lalicojournals.com

Volume 3; Issue 3; December 2025, Page No. 98-108.

Deepfakes as Multimodal Semiotic Signs: Iconic, Indexical, and Symbolic Dimensions in Nigerian Social Media

*¹Mahmud, M.O., ²Ogunbodede, K. F. & Ewata, T.O.³

¹Department of English Studies, Adekunle Ajasin University, Akungba-Akoko, Ondo State, Nigeria

²Federal Polytechnic Orogun, Delta State, Nigeria

³Department of Linguistics, Foreign and Nigerian Languages, University of Africa, Toru-Orua, Bayelsa State, Nigeria

*Corresponding author email: yemimahmud@gmail.com

Abstract

This study examines deepfakes as multimodal semiotic signs operating simultaneously through iconic, indexical, and symbolic dimensions in Nigerian social media. Employing Peirce's triadic theory, it investigates how deepfakes generate meaning and erode public trust by exploiting fundamental processes of human meaning-making. Through semiotic content analysis of 20 deepfake samples circulated on WhatsApp and Facebook during Nigeria's 2023 general elections, combined with recognition testing among 420 participants in Ondo and Bayelsa States, this study demonstrates the stratified semiotic sophistication that distinguishes deepfakes from conventional misinformation. Results reveal remarkably high communicative effectiveness across all three semiotic dimensions: 82.3% accuracy in iconic recognition (visual and auditory similarity to target personalities), 76.1% success in indexical interpretation (understanding fabricated character associations and causal relationships), and 71.8% accuracy in symbolic comprehension (decoding culturally encoded meanings, ethnic symbolism, and political narratives). This multimodal semiotic operation creates layered plausibility structures that prove extremely resistant to detection and critical evaluation. The findings document severe trust erosion in Nigeria's digital information ecosystem, with only 18% of respondents expressing trust in social media information, while 76% reported declining trust over the past year. Paradoxically, interpersonal trust (67%) remains significantly higher than institutional trust (8% for WhatsApp groups, 11% for Facebook), creating vulnerabilities as deepfakes circulate through trusted social networks. The study reveals how deepfakes exploit visual evidence bias, transform social capital into misinformation transmission channels, and generate persistent reputational damage even after exposure - the "liar's dividend" phenomenon. These semiotic operations pose acute threats to democratic processes, with 94% of participants perceiving deepfakes as threatening fair elections in Nigeria's multilingual, culturally diverse information landscapes.

Keywords: Deepfakes; Multimodal Signs; Nigerian Elections; Peircean Semiotics; Public Trust

Introduction

Communication has developed since the early oral traditions developed by use of electronic media until the present digital media, with each era presenting strategies that alter the way the human being crafting and conceiving meaning. Social media platforms in Nigeria today are critical arenas in the society where politics is debated and identity as a group is created. The deepfake technology, which consists of the use of advanced AI algorithms creating believable manipulations, is both a technical breakthrough and an acute communicative challenge with serious threats to the integrity of information, democratic processes, and even social solidarity. Social media platforms are the main sources of information in the African most populous country, Nigeria that has more than 200 million inhabitants and 154 million internet users. Nigerian digital space is distinguished by the outstanding linguistic diversity amounting to more than 500 languages, which generates complicated communicative spaces where information moves in more than one direction. This ambiguity generates vulnerabilities that the deepfaker designers take advantage of to influence the general opinion. The 2023 Nigerian general elections were a watershed in terms of the use of fabricated videos where politicians make false claims, which were popular across the country. Other high-profile examples were an alleged audio tape of a conversation between the presidential candidate of the Labour Party Peter Obi and Bishop David Oyedepo, which received more than 10 million views before being declared a deepfake. The deepfakes are semiotic signs that have complex meanings that work on a

number of planes. The triadic model of Charles Sanders Peirce with the three categories of iconic signs (depending on resemblance), indexical signs (depending on causality), and symbolic signs (depending on cultural associations) gives a solid model on how deepfakes generate communicative outcomes. On the iconic level, deepfakes take advantage of visual and audio similarity to real persons. On the indexical level, they build manifested causal relations between the persons being represented and the acts depicted. On the symbolic level, they resort to the common cultural knowledge and social stereotypes to make it more believable. Such a level of semiotic sophistication is what separates deepfakes with other forms of less advanced misinformation and why they are incredibly effective. The spread of deepfakes has caused a massive crisis of societal credibility in online information. Repeated exposure to manipulated material leads the audiences to adopt maladaptive coping mechanisms and they will become excessively credulous or excessively cynical. The two answers are a disincentive to informed democracy. Although the concept of deepfakes is becoming a subject of scholarly interest worldwide, the research on the issue that incorporates semiotic analysis and simultaneously focuses on the impact measurement is especially scarce in the context of Nigeria. The research fills this gap by using the Peircean semiotics to examine how deepfakes function as signs in terms of iconic, indexical, and symbolic scales and evaluate their effect on the general trust of the population of the social media users in Nigeria.

Theoretical Framework

This paper uses the triadic theory of signs developed by Charles Sanders Peirce to interpret deepfakes as the communicative phenomenon. Peirce described a sign as something that is to somebody something in some respect or capacity, and semiotics can be defined as the concept of the three inseparable parts that are representamen or sign itself, the object or what is being referred to, and interpretant or the meaning that is formed in the mind of the interpreter. This three-relationship explains such intricate interactions between the products of synthetic media, real people being modeled, and meanings created in the minds of viewers. The analytical structure is the classification of signs in three basic categories offered by Peirce. The iconic signs work based on similarity to the objects. Deepfakes involve using artificial faces, voices and motions created by AI algorithms that are highly realistic and similar to real persons due to Generative Adversarial Networks that are trained on large data sets that allow pixel-level similarity. This iconic correctness produces preliminary perceptual deceit. The indexical signs function in a direct relation or causality between sign and object. In deepfakes, when the audience watches a video in which a politician is delivering a statement, the audience perceives this as indexical evidence that the politician himself delivered the statement. The video seems to catalogue an actual situation, forming false assumptions concerning actions represented. This is especially strong as human beings have developed the reliability of visual and auditory indicators as true indicators of reality. Symbolic signs work on the basis of conventional or culturally acquired association. The symbolic dimension in the context of deepfakes aimed at Nigerian audiences is based on the culturally specific knowledge, social stereotypes, regional associations, and political stories. A deepfake of a politician wearing traditional clothing of a specific ethnic group may represent a symbolical elicitation of stereotyping concerning the nature of that ethnic group, and the symbolic connotation may rely on the cultural knowledge of the audiences with respect to the Nigerian ethnic dynamics. A critical theoretical point is that deepfakes affect all three dimensions of semiotic levels at the same time with a stratifying influence, which is extremely hard to resist and detect. This multimodal semiosis separates the deepfakes and simpler forms of misinformation that commonly work on single or two levels. Deepfakes successfully find a semiotic convergence by combining iconic similarity, indexical implication, and symbolic meaning in single communicative objects. This overlap brings about a layered plausibility whereby when the audiences start growing skeptical in one-dimension, other dimensions keep holding faith. Deepfakes also take advantage of the abductive reasoning, which can be described as the thinking process of deriving the most plausible explanation of the observed phenomena. Viewing deepfakes videos, the audience unconsciously engages in abductive inference, which is natural and automatic, but deepfakes would steal the process, and audiences will make false conclusions using otherwise valid processes of reasoning. The spread of deepfakes on a semiotic level essentially throws the indexical relation of signs and reality that forms the foundation of confidence in information into disarray. Once the audience cannot use visual and auditory evidence as credible measures of the real happenings, the whole structure of the evidential reasoning falls out of place, which brings indexical anxiety or a state of psychological insecurity regarding whether any particular sign can be considered credibly linked to the reality.

Methodology

This research was designed with a mixed-method approach using semiotic content analysis, recognition test and survey to evaluate. The study was carried out in six sites in Ondo State which are in Akure, Ondo Town, and Akungba-Akoko, and Bayelsa State which are in Yenagoa, Sagbama, and Toru-Orua. Among the 20 deepfake images and videos that were disseminated via WhatsApp and Facebook and reached considerable circulation during the 2023 general elections and the following 6 months in Nigeria, researchers found and archived these

images and videos. The key criteria were the content that was posted in at least 20 WhatsApp groups or with at least 10,000 Facebook interactions, and they were checked by the fact-checking organisations such as Africa Check and Dubawa. The sample consisted of 14 video deepfakes that were between 15 seconds and 2 minutes, 4 audio deepfakes, and 2 still images that focused on the high-profile Nigerian political leaders. The deepfake samples were of three presidential contenders of the 2023 elections in Nigeria representing three leading political parties in the country. In order to keep semiotic processes on the analytical track and safeguard individuals who fell prey to influenced media, applicants have pseudonyms where they are denoted as Candidate A, B and C during the presentation of the data. A case of Peter Obi and Bishop Oyedepo is widely disseminated in particular because of its unprecedented reach to the public and the fact that it has been documented to have an impact, and because it has been the subject of extensive analysis in Nigerian media and fact-checking articles. Each deepfake was analysed by two trained coders in categories of Peircean inquiry of iconic features such as facial similarity, voice, and contextual information, indexical features such as action represented and character assertion, and symbolic features such as cultural mention and ethnic affiliation. The inter-coder reliability was high with Cohen kappa of 0.78. Out of the total of the 20 deepfakes available in the archived collection, researchers chose 12 samples to conduct recognition tests. These 12 samples were randomly placed within the three semiotic dimensions and 4 samples were chosen directly to determine iconic recognition, 4 different samples chosen expressly to determine indexical interpretation and 4 different samples chosen expressly to determine symbolic understanding. This structure has made sure that each of the semantic dimensions is analysed separately using unique deepfake samples and that there is minimal testing effects and cross-contamination effects among types of assessments. The 4 samples to be used in testing iconic recognition were aimed at assessing how participants could recognise depicted individuals using visual and auditory similarity. The 4 samples to test the indexical interpretation were concerned with the capacity of the participants to interpret what was being said, done or what character traits were being ascribed on the people depicted. The 4 samples that aimed at testing the symbolic comprehension were concerned with how well the participants identified culturally encoded meanings, ethnic symbolism, and political versions. Any text, captions, and context information that would accompany the samples were taken out so that the participants would only use audiovisual information to draw their interpretations.

The research involved 420 respondents, purposely sampled and one hundred and ten in each of the Ondo State and Bayelsa State. The study was done in three universities in each of the 6 sites, and in community centres. The participant's age was between 18 years old to 65 years old with a mean age of 33.8 years with 52 percent being female and 48 percent being male. The level of education was 35 percent secondary education, 44 percent undergraduate education and 21 percent postgraduate education. Yoruba was 28 percent, Ijaw 26 percent, Igbo 19 percent and others were 27 percent. All participants actively used social media, daily using WhatsApp, Facebook or both and 91% of them said that they had encountered possible misinformation or fake content within the past month. There were inclusion criteria of being a citizen of Nigeria, aged 18 years or above, and a frequent user of social media and an informed consent was required. The testing was done in quiet and controlled conditions and protocols were done individually on tablets under supervision of researcher. The participants were first taken through a recognition testing protocol after giving informed consent in writing and filling demographic questionnaires. All the 12 deepfake samples were tested to each participant, and the testing was organised in such a way that it evaluated the three semiotic dimensions individually. The samples were randomised in order to eliminate sequence effects and the participants were not aware of what semiotic dimension would be evaluated in a given item. The design of the testing provided 1,680 responses per dimension (420 participants x 4 items per dimension) and 5,040 responses altogether across all three dimensions (420 participants x 12 total items). The test was conducted within three phases that are associated with the three dimensions of semiotics. During the first stage of the experiment that measured iconic recognition, the subjects were shown 4 deepfake samples and forced to determine who the individual in the sample was or who the individuals in the sample were by choosing between multiple-choice alternatives with the actual target and three distractor personalities of the same prominence. When the participants were able to recognise the target individual, this was rated as success. During the second stage of indexical interpretation, the participants observed 4 various deepfake samples and were requested to recognise what action, statement, or character trait was being projected onto the person in the picture by choosing one of the multiple-choice variants reflecting the possible interpretation. The success was rated in cases where the participants identified the characterisation in which the creator wanted to be identified correctly. The third stage with symbolic comprehension was used whereby the participants were exposed to 4 dissimilar samples of deepfakes and were requested to recognise the cultural, political, or social meaning in the content by choosing between multiple-choice alternatives portraying varying symbolic meanings. Success was recorded whereby the participants were able to identify the intended symbolic message correctly. All of the test items had four response options, one of which was the correct answer that represented the intent of the creator of the test as identified after the initial semiotic content analysis and confirmation with fact-checking organisations. After the recognition test,

the respondents were asked to answer a 25-item Likert-scale survey of trust in the information provided by social media, their perceived prevalence of deep fakes, the extent to which they believed that deep fakes affected prominent figures, and the extent to which they believed they affected the democratic process. The survey has taken a 5-point scale with a range of 1 (Strongly Disagree) to 5 (Strongly Agree) and individual items and composite scales were used to measure various areas of trust erosion and democratic influence. Some respondents also gave open-ended answers to the qualitative questions concerning their experience with deepfakes, verification methods, and advice on how to solve the issue, and 200 respondents answered this qualitative part. The research ethics committees of Adekunle Ajasin University and University of Africa gave their consent to the study. Detailed information about the purpose of the study, procedures, risks involved, and their right to withdraw at any time without penalties was given and all participants gave written informed consent. After all the testing and surveys, the participants were debriefed with adequate disclosure of the fact that some of the content they viewed was manipulated, information on the deepfake technology and how it can be detected, and access to fact-checking and media literacy. The participants who conveyed their distress in terms of the implication of deepfake technology were given more support and counselling materials.

Data Analysis

The data of recognition testing were processed with the help of SPSS Version 27.0 with the purpose to compute the success rates of each semiotic dimension. Chi-square tests were used to compare the relationship between demographic variables of a nature that was categorical and recognition success, independent samples t-tests compared the performance across states. One-way ANOVA involved examination of differences between groups of age, levels of education, and ethnic groups. The multiple regression analysis tested concomitant influences of demographic factors, previous awareness of deepfakes, belief in the power to detect it, and the frequency of social media use on general scores of comprehensions. The analysis of survey data was based on descriptive statistics of means, standard deviations, and percentage of agreement. Cronbach Alpha Coefficients were applied in determining internal consistency reliability. Mean scores were compared using independent samples t-tests between state scores and one way ANOVA compared scores between demographic groups. Thematic analysis of 200 qualitative responses where participants were asked open-ended questions was carried out in line with the 6-step method mentioned by Braun and Clarke. Independent coding of answers was done by two researchers, and the inter-rater reliability of the participants was found to be at Cohen level of kappa of 0.76. Chi-square tests were conducted to determine the difference in the frequency at which the themes varied across states. Semiotic content analysis entailed logical coding of two trained coders using operationalised definitions of each category in the Peircean classification. The results of inter-coder reliability had coefficients of Cohen kappa of 0.78 with percentages and frequencies, based on each coded feature, being computed.

For state-level comparisons, overall comprehension scores were calculated as the mean of the three-dimensional scores (iconic, indexical, and symbolic) for each participant, then averaged across all participants within each state.

Results and Analysis

The semiotic analysis of 20 deepfake samples showed the advanced usage in a variety of dimensions, as presented in Table 1. Platform distribution reflected a 60% circulation in the form of WhatsApp private groups and 40% reflected a viral spread on Facebook. The distribution of the deepfakes targeted Candidate A with corruption and other financial misconduct themes (7), Candidate B with ethnic bias and other inflammatory statement themes (8), and Candidate C with the religious insensitivity themes (5). The high success of the video format (70% of samples) over the audio (20% of samples) and the non-imageable format (10% of samples) is indicative of the current state and potential of deepfake technology and the desire of creators to employ multimodal content with visual and audio to have a stronger persuasive impact.

Table 1: Deepfake Sample Characteristics and Semiotic Features (N=20)

Characteristic	n	%	Semiotic Quality
Platform Distribution			
WhatsApp	12	60	Private group circulation, interpersonal trust
Facebook	8	40	Public viral spread, algorithm amplification
Content Type			
Video	14	70	Multimodal audiovisual integration
Audio	4	20	Voice cloning technology
Static Image	2	10	Facial manipulation only

Characteristic	n	%	Semiotic Quality
Target Distribution			
Candidate A	7	35	Corruption/financial misconduct themes
Candidate B	8	40	Ethnic bias/inflammatory statements
Candidate C	5	25	Religious insensitivity themes
Iconic Fidelity			
High facial resemblance (>80%)	16	80	Advanced AI facial mapping
Voice mimicry present	12	60	Sophisticated voice cloning
Contextual authenticity	16	80	Authentic settings/backgrounds
Indexical Features			
Corruption/financial misconduct	8	40	Accepting bribes, misappropriation
Ethnic bias/hate speech	6	30	Inflammatory ethnic statements
Incompetence	4	20	Fabricated gaffes, confusion
Religious insensitivity	2	10	Sectarian statements
Symbolic Features			
Ethnic symbolism	13	65	Traditional attire, language choices
Political narratives	11	55	Campaign themes, unity threats
Regional associations	9	45	Geographic references, rivalries
Class/wealth symbolism	7	35	Ostentatious displays, poverty
Religious symbolism	6	30	Religious language, settings

Note: Symbolic features are not mutually exclusive; individual deepfake samples could contain multiple symbolic dimensions. Percentages represent the proportion of samples containing each feature and therefore sum to more than 100%. Platform distribution reflects primary circulation channel. All samples were verified by Africa Check or Dubawa fact-checking organisations.

The iconic dimension analysis showed that 16 out of 20 samples recorded a high level of resemblance of over 80% fidelity in facial features, tone of skin and unique features of target personalities. Voice mimicry existed in 12 samples, and it showed the advanced voice cloning technology that imitated speech patterns, accents, and intonations of emotions. Contextual authenticity was present in 80 percent of the 16 samples, and creators were keen to select or even make backgrounds like campaign offices, government buildings or cultural facilities to make them look authentic. A single example was a 38-second video that was targeting a presidential candidate, which achieved a 90 percent overlap with the candidate's face, due to the proper mapping of unique facial features and skin tone, and to voice synthesis, recreating the typical patterns of Igbo-accented English word pronunciation, and shot in what seemed to be the campaign office, with recognisable campaign posters in the background. The indexical dimension analysis revealed that 8 samples at 40% represented corruption or financial misconduct such as taking bribes or embezzling funds, 6 samples at 30% represented ethnic bias or hate speech such as inflammatory statements about specific ethnic groups, 4 samples at 20% represented incompetence such as fabricated gaffes or confused statements, and 2 samples at 10% represented religious insensitivity. These indexical characteristics worked in the establishment of seemingly causal relationships between targets and morally objectionable behaviors that made viewers deduce defects in characters based on displayed behaviours. The deepfakes whose targets were corruption centred often took the form of compromising events, like receiving cash envelopes or having suspicious interactions about money, and the Nigerian historical issues revolving around political corruption were used to add credibility to the alleged deepfakes. The deepfakes involved in the ethnic bias portrayed the targets to make sectarian proclamations that are based on the ethnic relations and past conflicts in Nigeria to imply prejudice and rifts. The symbolic dimension analysis showed that 13 samples out of 65 percent used ethnic symbolism in terms of using traditional clothes, use of language or regional setting that elicited certain ethnic group connotations and stereotypes. There were 55 percent instances of political stories in 11 samples, which integrated content in larger campaign stories like having certain candidates being perceived as a threat to national unity or regional interests. In 9 samples, regional associations were present at 45 percent, with geographical references being strategically used to appeal to North-South or interstate level to invoke rivalries. The symbolic use of class and wealth was presented in 7 samples with 35 percent and portrayed the targets in glitzy wealth showcases to indicate corruption, or in poor environments to pose a challenge to their efficiency.

Religious symbolism was found in 6 samples at 30% where religious language, setting, or characters were called on to invoke sectarian tensions. Such symbolic aspects demanded that audiences had to have some cultural information regarding how the Nigerian ethnicity, political history, and its religious populations, and social stereotypes in order to understand the intended meanings.

Table 2: Recognition Testing Results Across Semiotic Dimensions (N=420)

Semiotic Dimension	Success Rate (%)	Correct/ Total Responses	Range Across Items	State Comparison	Education Comparison	Statistical Significance
Iconic Recognition	82.3	1,384/1,680	79%-86%	Ondo: 83.2%; Bayelsa: 81.4%	Secondary: 80.4% Undergraduate: 82.3% Postgraduate: 84.6%	All dimensions above chance (50%) State: $t(418)=3.87$, $p<.001$
Indexical Interpretation	76.1	1,278/1,680	73%-79%	Ondo: 77.1%; Bayelsa: 75.2%	Secondary: 73.8% Undergraduate: 75.1% Postgraduate: 79.4%	Education: $F(2,417)=11.23$, $p<.001$
Symbolic Comprehension	71.8	1,206/1,680	68%-75%	Ondo: 74.2%; Bayelsa: 69.5%	Secondary: 69.2% Undergraduate: 70.8% Postgraduate: 76.1%	Ethnicity on symbolic: $F(3,416)=8.94$, $p<.001$ Yoruba: 76.8%* vs Others: ~69%
Overall Comprehension	76.7	3,868/5,040	-	Ondo: 77.6%; Bayelsa: 75.3%	Secondary: 74.5% Undergraduate: 76.1% Postgraduate: 79.4%	Age: $F(2,417)=7.89$, $p<.001$; 18-29: 74.4% vs 46-65: 78.7%

Note: All success rates significantly exceeded chance level (50%) at $p<.001$. *Indicates Yoruba participants significantly outperformed other ethnic groups on symbolic comprehension at $p<.001$.

The communicative effectiveness of deepfakes was remarkably high across all semiotic dimensions in recognition testing. Iconic recognition had an 82.3% success rate with 1,384 correct identifications out of 1,680 total responses, with individual items ranging from 79% to 86% accuracy. The participants proved to be very effective in recognising target personalities using visual and auditory similarity only, which demonstrates the iconic sophistication of the analysed deepfakes. The success rate of indexical interpretation was 76.1%, with 1,278 correct interpretations out of 1,680 responses, ranging from 73% to 79% across the items. Participants were able to decode intended character associations and moral judgments implicit in the actions represented, which showed that indexical dimensions were effective in conveying creators' intended messages concerning the personality and actions of targets. Symbolic comprehension had a 71.8% success rate with 1,206 correct interpretations among 1,680 responses, ranging from 68% to 75% by item. The demographic analysis showed that there were significant differences in comprehension patterns. Age was a positive predictor of comprehension, with the 46 to 65 age group achieving 78.7% mean accuracy compared to the 18 to 29 age group which had 74.4% accuracy, $F(2, 417) = 7.89$, $p < .001$. This age difference was more apparent in symbolic comprehension, where older participants achieved higher accuracy rates (75.2%) compared to younger participants (67.8%), which could be attributed to greater familiarity with Nigerian political histories, ethnic tensions, and cultural allusions to which the symbolic dimensions appealed. Performance was strongly predicted by education level, with postgraduate participants scoring 80.0% average accuracy versus secondary education participants at 74.5%, representing a highly significant difference, $F(2, 417) = 11.23$, $p < .001$. This education effect was consistent across all three dimensions, indicating that formal education enhances perceptual discrimination skills for iconic recognition as well as critical thinking skills for indexical and symbolic interpretation. There were no significant gender differences, with male and female participants performing comparably across all dimensions. Ethnicity had a major impact on symbolic understanding, $F(3, 416) = 8.94$, $p < .001$, with Yoruba participants demonstrating 76.8% accuracy versus 69.3% for Ijaw participants, 67.4% for Igbo participants, and 68.9% for other ethnicities. This pattern probably indicates that many of the symbolic elements in the deepfake samples were more aligned with South-West Nigerian (Yoruba) cultural contexts than South-South or South-East ones. The comparison between participants from Ondo

State and Bayelsa State revealed that the former achieved an overall comprehension rate of 77.6% versus 75.3% for the latter, a statistically significant difference, $t(418) = 3.87, p < .001$. The regression results which were determined as significant predictors of overall comprehension as seen in Table 3. Education level presented standardised beta of 0.28 which indicated positive strong relationship, previous deepfake awareness presented standardised beta of 0.31 that indicated strongest predictor, confidence in detection ability presented standardised beta of 0.24, and frequency of using social media presented standardised beta of 0.19. The combination of these variables contributed to 38.7% of variance in comprehension scores $F(7, 412) = 37.24, p < .001$. The independent contributions of age, gender, and ethnicity were insignificant, which indicated that the relationships between these variables were influenced by differences in education, awareness, and digital literacy instead of demographic influences. The conclusion that the experience of deepfake awareness beforehand appeared as the most powerful might tell significant implications in the intervention measures, and educational programs that aim at creating awareness regarding the use of deepfake technology and its peculiarities may help to develop the detection skills of the target population significantly.

Table 3: Multiple Regression Analysis Predicting Overall Comprehension (N=420)

Predictor Variable	B	SE	β	t	p	95% CI
Education level	3.42	0.68	.28	5.03	<.001	[2.09, 4.75]
Prior deepfake awareness	4.18	0.72	.31	5.81	<.001	[2.77, 5.59]
Confidence in detection	2.89	0.64	.24	4.52	<.001	[1.64, 4.14]
Social media use frequency	2.15	0.69	.19	3.12	.002	[0.80, 3.50]
Age	0.87	0.58	.08	1.50	.135	[-0.27, 2.01]
Gender	-0.34	0.52	-.03	-0.65	.516	[-1.36, 0.68]
Ethnicity	0.61	0.48	.06	1.27	.205	[-0.33, 1.55]

Note: $R^2 = .387, Adjusted R^2 = .376, F(7, 412) = 37.24, p < .001. \beta = standardised\ beta\ coefficient.$

Table 4 describes the result of a survey that indicated the lack of trust in the information provided in the social media. The general trust on social media was found to be 2.81 out of 5 with a standard deviation of 0.73 which was far below the midway mark of 3 whereby only 18% of respondents indicated that they believed that information found in social media is trustworthy or strongly trustworthy. WhatsApp group information was the lowest at 2.18 where only 8% stated that they trusted the information whilst Facebook content scored 2.34 where 11 percent said they trusted. Nonetheless, the level of trust regarding information posted by friends and relatives was comparatively better at 3.45 whereby 67 percent of the respondents mentioned that they trusted people, which indicates that interpersonal trust does not go away when institutional trust becomes weak. Interestingly, 59 percent of the respondents said they trusted videos and images more than text-based content, which proves that there is always a psychological tendency to trust visual and auditory data, as deepfakes take advantage of this fact and rely on the evidentiary value that human beings tend to place on visual and aural data. Worst of all, 76 percent have become less trusting of the information they receive on social media in the last year, recording a rapid decline in their trust as deepfakes and fake news become more common and more advanced.

Table 4: Trust Perceptions and Democratic Impact (N=420)

Measure	M	SD	% Agree/ Strongly Agree	Key Finding
Trust in Social Media Information				
Overall trust in social media	2.81	0.73	18%	Severe institutional trust erosion
Trust WhatsApp group information	2.18	0.85	8%	Lowest trust score
Trust Facebook content	2.34	0.79	11%	Institutional skepticism
Trust friend/family shares	3.45	0.91	67%	Interpersonal trust persists
Trust videos/images more than text	3.52	1.01	59%	Visual evidence bias
Trust decreased over past year	3.89	0.95	76%	Accelerating decline
Perceived Deepfake Prevalence				
Personally encountered deepfakes	4.06	0.77	86%	Universal recognition
Deepfakes becoming harder to detect	4.01	0.83	81%	Technology sophistication
Fake news spreads faster than real news	4.15	0.74	89%	Viral misinformation

Measure	M	SD	% Agree/ Strongly Agree	Key Finding
Most users lack detection skills	4.09	0.79	84%	Collective vulnerability
Impact on Public Figures	4.23	0.58	91%	Widespread concern
Deepfakes damage reputations unfairly	4.18	0.65	90%	Unjust harm
Build negative stereotypes	4.21	0.63	91%	Stereotype reinforcement
Used for character defamation	4.29	0.59	94%	Malicious deployment
Damage persists after exposure	4.17	0.68	89%	"Liar's dividend" effect
Impact on Democratic Processes	4.31	0.54	94%	Democratic threat
Threaten fair elections	4.28	0.62	93%	Electoral integrity
Make informed voting difficult	4.19	0.67	89%	Voter confusion
Influenced 2023 elections	4.05	0.79	78%	Historical impact
Political opponents weaponise them	4.35	0.58	96%	Normalised tactic
Undermine democratic institutions	4.11	0.71	86%	Institutional erosion

Note: Scale: 1 = Strongly Disagree, 5 = Strongly Agree. "Agree/Strongly Agree" represents responses of 4 or 5 on the scale. All composite scales (bolded headings) showed high internal consistency with Cronbach's $\alpha > .84$.

There was a high score on the perceived prevalence of deepfakes of 4.06 out of 5 and standard deviation of 0.77 with 86 percent of the respondents agreeing or strongly agreeing that deepfakes are prevalent on the Nigerian social media platforms. This is particularly because 73% of them have themselves encountered suspected deepfakes, 81 percent believed that deepfakes are becoming more difficult to spot with advancing technology, 89 percent agreed that fake news travels more quickly than real news on social media, and 84 percent considered that the majority of social media users lack the ability to spot deepfakes. Influence on public figures was highly congruent with an agreement score of 4.23 of 5, and standard deviation of 0.58 with 91 percent of the respondents agreeing that deepfakes harm reputations, and develop negative stereotypes. Even 94% of them stated that deepfakes are applied to defame the character and 89% stated that the damage is still caused even when the deepfakes are reported deceitful, which is the dividend phenomenon of a liar. The most important one was impact on democratic processes with a standard deviation of 0.54 and 94% agreeing that deepfakes threaten fair elections, 89% agreeing that they made it hard to vote smartly, 84% agreeing that deepfakes were used to impact the results of the 2023 election in Nigeria, 96% agreeing that political opponents use deepfakes to attack one another and 86% agreeing that deepfakes compromise democratic institutions. The qualitative interpretation of 200 open-ended responses showed that 59 percent reported experiencing political content on the 2023 elections, 35 percent reported being a victim of commercial scam, and 24 percent reported being a victim of religious content. Checking several sources was a verification strategy that was used by 65 per cent, social validation with trusted persons was used by 48 per cent and fact-checking websites were used by only 23 per cent. Technological accessibility, educational deficits, political motivations, and cultural trust pattern were found to be contributing factors with 80, 71, 64, and 55 respectively. In Bayelsa, the participants were much more likely to mention linguistic complexity of 55% as opposed to 28% in Ondo. The solutions that were recommended were educational initiatives, 85%: government regulation, 65% technology solutions, 59%. A good example was the audio deepfake of Peter Obi and Bishop Oyedepo that went viral. In this deepfake, 90 percent of voice replication was reproducible with 87 percent of participants recognising both speakers. Indexically, 82 percent understood fabricated conversation that there was a religious bias. Symbolically, 68 percent asked the religious-political tension invocation. The deepfake has spread across more than 200 WhatsApp groups with 10 million views and 73% of those who saw it said they lost confidence in Obi and 61% said they still doubted after debunking.

Discussion

The results indicate that deepfakes function as highly effective multimodal semiotic signs that successfully convey intended meanings through the simultaneous operation of iconic, indexical, and symbolic dimensions. The high success rates in recognition testing—82.3% for iconic recognition, 76.1% for indexical interpretation, and 71.8% for symbolic understanding—provide evidence that deepfakes effectively exploit core processes of human meaning-making as conceptualised in Peircean semiotic theory. The effectiveness of the iconic dimension is indicative of the powers of the deepfake technology to create hyper-realistic content that cannot be differentiated by human perception systems with a real recording. The fact that 59 percent of the participants believed videos and photos more than text proves that deepfakes take advantage of psychological biases in which humans give high evidentiary value to visual and auditory data. The effectiveness of the indexical dimension proves that

deepfakes are effective in causing false beliefs concerning causal relationships between targets and actions described. This plays on abductive logic in which the observers subconsciously deduce that the visual evidence translates to events having taken place. The high efficiency of the symbolic dimension, even though it needs knowledge that is culturally particular, proves that deepfakes are effective in terms of triggering shared cultural schemas. The fact that Yoruba participants did better in symbolic comprehension by 7 percentage points than the other ethnic groups implies that symbolic meanings are culturally particular, and their usefulness relies on the presence of relevant cultural knowledge by the audiences. It is the combination of all three semiotic dimensions that produces stratified plausibility structures that are very hard to spot and hard to overcome. Even those who doubt one of the dimensions are skeptical and find other dimensions still supporting belief. Such multimodal semiosis is a qualitative amplification of the level of fakery information over the textual fake news that functions at a lower number of semiotic levels. The reported loss of trust, where the social media trust is at 2.81 of 5 in general and 76 percent of people experienced a loss of trust in the last one year is a deep crisis of the information ecosystem in Nigeria. The fact that 67 percent still trusts the content shared by friends and family and only 8 percent trust WhatsApp groups is a kind of paradox of the personal trust existing even as institutional trust fails, which makes the interpersonal trust vulnerable as the deepfakes that spread into personal networks use social capital. This tendency is indicative of what Bourdieu (1986) itself imaginatively termed as the transformation of the social capital of trusted relations into the means of misinformation transmission. By spreading deepfakes in friend networks or family WhatsApp chats, deepfakes also acquire the credibility of the relationships, and the recipient is less likely to engage in critical analysis. This is especially hazardous in the collectivist cultural environment of Nigeria where communal affiliation and individual acquaintances are conventionally the main information confirmation systems. This scam of these trust networks is a very basic corruption of a social capital which turns the best source of communicative power by communities into their greatest weakness.

The fact that 89% of people feel that reputation damage is equally enduring even after the deepfakes are detected is an example of liar dividend phenomenon in which deepfake technology introduces an ongoing uncertainty. The fact that it is now possible to convincingly create fake videos means that audiences can never be fully able to trust a certain video after checking its facts because they know that it is a perfect fake. This gives a bad actor an asymmetric advantage as they implement deepfakes based on the knowledge that even unsuccessful deception instills a sense of uncertainty in the long run. The prevalence of views that deepfakes are a threat to democratic processes, 94 percent of whom believe it is a threat to fair elections, and 96 percent of whom believe it is used as a weapon by political opponents, show that deepfakes have become a normal political instrument. To the extent that 84 percent thought that deepfakes affected the 2023 elections, it would imply that even though the elections may not have been affected by the fake, the belief that it was caused by them eroded trust in the legitimacy of democracy. The regression analysis results, especially the strongest effect of the previous awareness with the beta of 0.31, indicate that the detection abilities can be improved through the educational interventions. The level of education ($b=0.28$), belief in detection ability ($b=0.24$), and frequency of social media use ($b=0.19$) were all positively correlated with comprehension and these four variables collectively accounted to 38.7% of the variance at $F(7, 412) = 37.24, p < .001$. This gives the media literacy programs empirical support as important elements of holistic response programs. It is especially promising that the awareness of deepfakes precedence proved to be the unquestionably most significant predictor, which can be easily and systematically raised with the help of educational programs, unlike such demographic variables as age or ethnicity, which are difficult to alter. Nonetheless, the explanation of the model (only 38.7%) suggests that there are other unexplained factors, which may include cognitive abilities, critical thinking dispositions, or technological familiarity among others, that contribute significantly to deepfake understanding. These other predictors need to be studied in future to formulate more holistic education interventions.

Conclusion

This paper has shown that deepfakes are highly multimodal signs of semiotic meaning that significantly achieve the intended communicative effect by combining iconic, indexical, and symbolic layers of meaning in their operation and operate remarkably effectively, which is why these images have been reported to erode popular trust in Nigeria digital information ecosystem. The fact that the recognition testing rates are very high proves that deepfakes effectively take advantage of the basic processes of the human meaning-making, forming the levels of plausibility that are very hard to identify and resist. The universal loss of trust, as three-quarters of respondents indicated their reduction in trust in the last one year, is a deep-seated crisis to the process of democracy and social unity in Nigeria. The study shows that deepfakes have become an established instrument of politics in the Nigerian electoral arena, with 94% of participants perceiving them as threatening fair elections and 96% believing they are weaponized by political opponents. The liar's dividend effect, in which reputation damage persists even after deepfakes are exposed as false, gives malicious actors an asymmetric advantage and essentially renders conventional verification systems insufficient. Necessary reactions should combine several strategies. The media

literacy education ought to be carried out in a systematic manner in order to develop the ability of critical assessment and awareness of deepfake technology. Laws ought to be created to ban malicious deepfake production and safeguard free speech. New technological game changers such as AI-based recognising instruments must be created and made available. The fact-checking organisations are to be empowered by funding them higher and creating awareness among people. Interagency interventions between government agencies, technology firms, civil society organisations, educational institutions and media professionals are necessary. Since Nigeria is very diverse in terms of language, the intervention needs to be tailored to suit various environments since the deepfakes and verification methods may not work equally well across the various Nigeria communities. It will take a long-term intervention by various stakeholders to restore the confidence of the people in Nigeria in the digital information ecosystem. Although conditions seem to be hopeless at the moment, the fact that education-based interventions are capable of upgrading the detecting skills gives an impression that holistic actions can restore integrity of information. The case of Nigeria is a warning and call to action to have joint efforts before it is too late to save democratic institutions.

References

- Apuke, O. D., & Omar, B. (2020). Fake news and COVID 19: Modelling the predictors of fake news sharing among social media users. *Telematics and Informatics*, 56, Article 101475.
- Atkin, A. (2013). Peirce's theory of signs. In E. N. Zalta (Ed.), *The Stanford encyclopaedia of philosophy* (Summer 2013 ed.). Stanford University. <https://plato.stanford.edu/entries/peirce-semiotics/>
- Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). Greenwood Press.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Busari, S., & Adebayo, B. (2023). How deepfakes and AI are being weaponised in Nigeria's 2023 elections. African Digital Rights Network.
- Chandler, D. (2017). *Semiotics: The basics* (3rd ed.). Routledge.
- Chandler, D. (2020). Semiotics and AI: Towards a new semiotic paradigm. *Semiotica*, 234, 1–19.
- Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107(6), 1753–1820.
- Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2023). *Ethnologue: Languages of the world* (26th ed.). SIL International.
- Ekpong, P., Adeyemi, A., & Okon, E. (2023). Deepfake deployment in Nigeria's 2023 elections: Prevalence, motivations, and impacts. *Journal of African Elections*, 22(1), 45–68.
- Eleanya, C. (2023, March 12). Peter Obi audio: How deepfake conversation went viral. *The Guardian Nigeria*. <https://guardian.ng/news/peter-obi-audio-how-deepfake-conversation-went-viral/>
- Goodfellow, I., Pouget Abadie, J., Mirza, M., Xu, B., Warde Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (Vol. 27, pp. 2672–2680).
- Goudge, T. A. (1965). Peirce's index. *Transactions of the Charles S. Peirce Society*, 1(2), 52–70.
- Guanah, J. S., & Omorinola, A. (2018). Communication evolution in Nigeria: From oral traditions to digital platforms. *African Communication Research*, 11(2), 123–145.
- Harsin, J. (2015). Regimes of post truth, post politics, and attention economies. *Communication, Culture & Critique*, 8(2), 327–333.
- Mi, L., & Apuke, O. D. (2024). Digital literacy and misinformation susceptibility in Nigeria. *Information Development*, 40(1), 89–107.
- Nigerian Communications Commission. (2023). *Industry statistics: Internet subscriptions*. <https://www.ncc.gov.ng/statistics-reports>
- Onigbinde, F., & Oloyede, S. (2024). Language barriers and fact checking challenges in multilingual Nigeria. *Language Policy*, 23(1), 67–85.
- Paavola, S. (2004). Abduction through grammar, critic, and methodetic. *Transactions of the Charles S. Peirce Society*, 40(2), 245–270.
- Peirce, C. S. (1931–1958). *Collected papers of Charles Sanders Peirce* (Vols. 1–8). Harvard University Press.
- Rini, R. (2020). Deepfakes and the epistemic backstop. *Philosophers' Imprint*, 20(24), 1–16.
- Short, T. L. (2007). *Peirce's theory of signs*. Cambridge University Press.
- Sonesson, G. (1998). The concept of text in cultural semiotics. *Sign Systems Studies*, 26, 83–114.
- Statista. (2023). *Number of internet users in Nigeria from 2018 to 2028*. <https://www.statista.com/statistics/183849/internet-users-nigeria/>

- Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining fake news: A typology of scholarly definitions. *Digital Journalism*, 6(2), 137–153.
- Tolosana, R., Vera Rodriguez, R., Fierrez, J., Morales, A., & Ortega Garcia, J. (2020). Deepfakes and beyond: A survey of face manipulation and fake detection. *Information Fusion*, 64, 131–148.
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 1–13.
- Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11), 39–52.