

How Can a Party Show Authorship of a Social Media Post or Other Electronic Communication?

Suppose that a defendant is charged with possessing fentanyl with the intent to sell it. The state's evidence includes a Facebook direct message, purportedly from the defendant to an informant, saying "just got some [China Girl](#), you want any?" An officer took a photograph of the direct message as it appeared on the informant's smartphone, and everyone is satisfied that the photograph fairly and accurately depicts the message. But the defendant objects to the introduction of the message on the grounds that there's no way to be sure that he wrote it. How might the state respond?

This is a question about authentication. Before evidence may be admitted, the proponent must show that it is authentic – that it "is what its proponent claims." [N.C. R. Evid. 901\(a\)](#). In many cases involving electronic communications, this is a two-step process. The first step is showing that the communication has been accurately captured. In our example, the officer took a photograph of the direct message, so there is no problem with this aspect of authentication.

The second step is showing that some relevant person, usually the defendant, is the author of the communication. That can be harder. As the Third Circuit put it in an often-quoted sentence, "the authentication of social media evidence in particular presents some special challenges because of the great ease with which a social media account may be falsified or a legitimate account may be accessed by an imposter." *United States v. Browne*, 834 F.3d 403 (3d Cir. 2016).

Authorship may be established through circumstantial evidence. Once in a while, there may be direct evidence of authorship: the purported author admits writing the communication, or someone testifies that he or she saw the purported author writing it. That's sufficient under N.C. R. Evid. 901(b)(1).

More often, there isn't direct evidence, and the proponent seeks to rely on circumstantial evidence of authorship, such as the communication coming from an account that is linked to the purported author, or the communication referring to events in which the purported author was involved. It is entirely proper to use this sort of circumstantial evidence to authenticate a communication. N.C. R. Evid. 904(b)(4) allows reliance on "distinctive characteristics" and "circumstances."

At this point, our state appellate courts and the federal circuit courts have decided quite a few cases on the authentication of electronic communications. What follows is a non-exclusive list of some of the types of evidence courts have found pertinent to authentication.

- Evidence that the communication came from an account previously used by the purported author, or from an account that is linked to such an account (e.g., from a Facebook account that is linked to an email address or phone number known to belong to the purported

- author). See, e.g., *State v. Allen*, 250 N.C. App. 823 (2016) (unpublished) (fact that message was sent from an account that defendant had used to communicate with others for years tended to show that the defendant was the author of the message).
- Evidence that the communication came from an account with a username that is, or includes, the purported author's name or nickname. See, e.g., *United States v. Banks*, 29 F.4th 168 (4th 2022) (fact that the defendant's name was the account username tended to show that he was the author of the account's contents); *United States v. Recio*, 884 F.3d 230 (4th Cir. 2018) (similar); *State v. Ford*, 245 N.C. App. 510 (2016) (fact that defendant's nickname resembled the name on a social media account tended to show that defendant was the account holder and author of posts on the account); *State v. Pryor*, 2021 WL 3354746 (N.C. Ct. App. Aug. 3, 2021) (unpublished) (fact that messages came from a Facebook account bearing the defendant's name was some evidence that the defendant was the author of the messages).
 - Evidence that the communication came from an account that contains other information associated with the purported author, such as pictures of the person and his or her friends and family. See, e.g., *United States v. Recio*, 884 F.3d 230 (4th Cir. 2018) (fact that social media account contained over 100 pictures of the defendant tended to show that he was the account holder and the author of the account's content); *State v. Ford*, 245 N.C. App. 510 (2016) (fact that social media page on which postings appeared also contained videos of the defendant's dog was evidence that the defendant was the account holder and the author of the account's content); *United States v. Hassan*, 742 F.3d 104 (4th Cir. 2014) (fact that communication came from an account that contained accurate biographical details about the defendant tended to show that the defendant was the account holder and the author of the account's content).
 - Evidence that the communication came from a device that was seized from the purported author or from an account installed on such a device. See, e.g., *United States v. Banks*, 29 F.4th 168 (4th Cir. 2022) (fact that communication came from an account that was installed on a phone seized from the defendant tended to show that he was the account holder and author of the account's content); *State v. Wilkerson*, 223 N.C. App. 195 (2012) (fact that text message was sent from a number associated with a phone later seized from the defendant tended to show that the defendant was the author of the message).
 - Evidence that the communication referred to specific facts or events known only to the purported author, or to a small number of people including the purported author, or of special interest or significance to the purported author. See, e.g., *State v. Allen*, 250 N.C. App. 823 (2016) (unpublished) (a communication referenced a gift the defendant had previously given to the recipient of the communication; this tended to show that the defendant was the author of the communication).
 - Evidence that the communication was connected in time or content to events in which the purported author was involved. See, e.g., *State v. Clemons*, 274 N.C. App. 401 (2020) (Facebook comments were likely to have been authored by the defendant in part because they were made shortly after he was released from prison); *United States v. Davis*, 918 F.3d 397 (4th Cir. 2019) (communications were likely to have been authored by the

defendant where the communications negotiated a time and place for a drug transaction and the defendant subsequently appeared at that time and place).

- Evidence that the communication originated from an IP or physical address associated with the purported author. *United States v. Hassan*, 742 F.3d 104 (4th Cir. 2014) (fact that account was connected to an IP address associated with the defendant tended to show that the defendant was the account holder and the author of the account's content).

There is no hard and fast rule about which of these factors, or how many, are sufficient to authenticate a communication. Often several factors are considered together. For example, in *United States v. Perez*, 61 F.4th 623 (8th Cir. 2023), the court ruled that the prosecution adequately authenticated child pornography images and other material obtained via search warrant from a social media platform as belonging to the defendant where “(a) the user of [the account] referred to himself [using the defendant's first name], (b) the account was created using [the defendant's] Google email address, (c) the account was regularly accessed by the IP address associated with [the defendant's ISP] account, and (d) the account contained numerous personal photographs of [the defendant]”). However, authentication is a relatively low hurdle and the trend in the case law is towards finding circumstantial evidence sufficient to authenticate authorship.

Returning to the example at the beginning of this post, how might the state respond to the defendant's claim that there's no evidence he wrote the message in question? Lots of ways. It might show that the message came from an account bearing the defendant's name, or containing content pertinent to him. It might show that the informant had communicated with the defendant previously using that account. Or it might show that when the informant expressed interest in the fentanyl and set up a time to make the transaction, the defendant showed up at the meeting. None of these responses would be conclusive proof of authorship, but they might nonetheless be sufficient to allow the evidence to be admitted and considered by the jury.