IS IT POSSIBLE TO IDENTIFY A CRIMINAL BY VOICE ALONE?

Harriet M. J. Smith[1]*, Nikolas Pautz[1] and Katrin Mueller-Johnson[2]

[1] Department of Psychology, Nottingham Trent University, Nottingham, United Kingdom

[2] Faculty of Law, University of Oxford, Oxford, United Kingdom

YOUNG REVIEWER:

ETHAN
AGE: 9

EARWITNESS

A person who has overheard, but not seen, something that might help the police solve a crime.

Imagine you overhear someone talking about a robbery they just committed. You hear their voice, but you can not see them. A few weeks later, the police ask you to identify the person you overheard, from a voice lineup. You listen to different peoples' voices and try to pick out the criminal. Do you think you would be able to do it? Perhaps you think you would be able to identify the criminal because you can easily recognize the voices of your family and friends. In fact, recognizing a stranger's voice is difficult. There are many reasons you might struggle to remember a stranger's voice accurately. We will learn about how memory for voices works, and how it can help us predict whether an "earwitness" will correctly select the guilty person. We will also learn how the police can help an earwitnesses to pick the criminal out of the lineup.

WOULD YOU BE A GOOD EARWITNESS?

An **earwitness** is someone who hears information that is useful to the police. The information an earwitness provides can help to convict

criminals. Imagine that you are an earwitness. You are on a train, and you overhear a stranger talking to his friend on the phone. He is talking about a robbery he just committed. You cannot see the stranger's face because he is sitting behind you. However, you can clearly hear his voice. You notice that he speaks with an unusual accent. The phone conversation continues for around 10 min. You report the conversation to the police as soon as you get home.

Two weeks later, the police tell you that they have arrested a **suspect**. To make sure they have the right person, they ask you to come into the police station and listen to a voice lineup. You listen to recordings of 9 different voices and try to identify the person from the train. You are told that only one of the voice recordings is that of the suspect, and that all of the other voices belong to people who are **foils**. Each voice recording lasts around 1 min. Once you have listened to all 9 voices, you try to identify the person from the train. You find it difficult and you are not very confident in your decision because the voices sound quite similar to each other.

After you have selected someone, the police thank you for your time and you go home. A few weeks later, you find out that the person you selected will be sent to prison for the robbery.

HOW ACCURATE IS VOICE IDENTIFICATION?

Voice identification can be used as evidence in court to help convict criminals. Should we trust voice identification evidence? Well, psychological research has shown that earwitnesses are likely to select the wrong person from a voice lineup [1]. If the police have arrested an innocent person, the wrong person might be sent to prison, and the guilty person would never be punished. There are three reasons why voice identification is difficult. First, when we try to identify a person, we normally depend on the person's face, not their voice. Second, when we listen to someone speaking, we tend to concentrate more on what they are saying, rather than what their voice sounds like. Finally, memory does not work like a video camera. Just because you heard something previously does not mean you will be able to remember it clearly later [2].

HOW DOES MEMORY WORK?

Psychological researchers often explain memory as having three stages (Figure 1). **Encoding** happens when you create a memory (like when you overhear a person discussing a robbery). **Storage** is the period of time between encoding and when your memory is tested (such as the 2 weeks before you visit the police station). **Retrieval** is when your memory is tested (when you try to identify the person from the train).

Figure 1
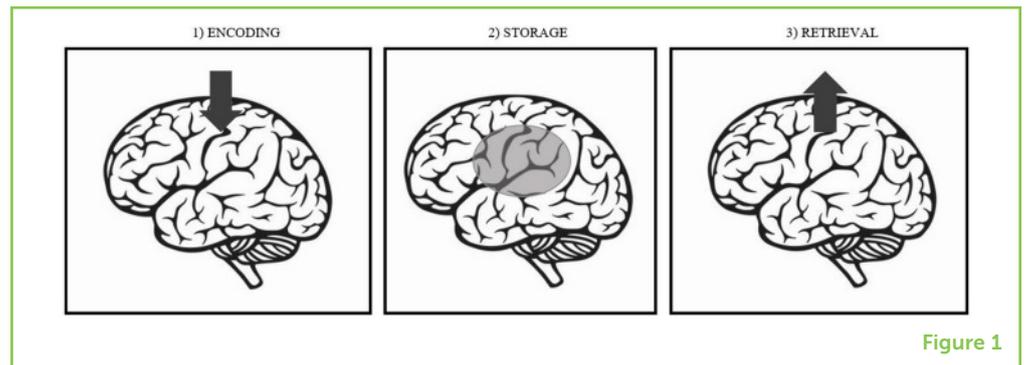
Memories are fragile and can be damaged at any of these three stages. Your memories are not necessarily accurate. Damage to the original memory can occur without you noticing. You may believe you have accurately identified someone from a voice lineup when in fact you have not. Equally, you may feel very unsure about your decision, but your decision might be correct.

Even if your memory for a voice is good, your voice lineup decision might be affected by your expectations. If you believe that the police are likely to have arrested the correct person, this might lead you to select one of the voices as the suspect. If you feel the police have arrested the wrong person, you might decide not to select anyone.

Based on the complex stages of memory, we know that earwitnesses are likely to select the wrong person from a voice lineup. To predict whether an earwitness has identified the correct person, we need to learn more about the ways memories can be damaged at the encoding, storage, and retrieval stages.

WHAT INFLUENCES VOICE IDENTIFICATION ACCURACY?

There are many variables that might affect the accuracy of voice identification. In earwitness research, these variables can be divided into two main categories: **estimator variables** and **system variables** [3].

Estimator Variables

ESTIMATOR VARIABLE

Something that might affect a witness' memory, but that **cannot** be controlled by the police.

SYSTEM VARIABLE

Something that might affect a witness' memory, that **can** be controlled by the police.

Estimator variables can damage memories at the encoding and storage stages. They are associated with the characteristics of the witness, the suspect, and the environment in which the event took place. Estimator variables have their name because the police were not at the crime, so they can only *estimate* the effect of these variables on memory. Understanding which estimator variables are relevant in a crime helps us to predict whether an earwitness is likely to be accurate.

However, once estimator variables have damaged a memory, there is nothing that can be done to undo the damage.

Several estimator variables may have been present during the conversation you overheard on the train. One is familiarity, which is how familiar you are with a person's voice. If you are familiar with a person's voice, you will probably be able to accurately identify that person in a voice lineup. If the voice belongs to someone you have never met, it might be more difficult [4]. On the train, you overheard a stranger's voice, so the voice lineup would have been particularly difficult.

Duration is another estimator variable. The amount of time that you spend listening to someone speaking can affect how difficult it is to pick their voice out of a lineup [5]. If you listen to a voice for a long time, you hear a greater variety of words and sounds. This means it might be easier to match these words and sounds to the suspect's voice in the voice lineup. On the train, you heard the stranger talking for around 10 min. This would make the voice lineup easier than if you had only heard the person saying a few sentences.

The last estimator variable is distinctiveness. Everyone's voice sounds different, but some voices are particularly distinctive, which means that they stand out from other voices. You are more likely to be able to identify a distinctive voice than an average-sounding voice [6]. The stranger on the train had an unusual accent. This may make the voice lineup easier for you.

System Variables

System variables relate to how the voice lineup is conducted. If the lineup is conducted badly, system variables can damage memory at the retrieval stage. On the other hand, system variables can be controlled by good police work to minimize the damage to memory.

Lineup procedure is one system variable. There are various ways of asking someone to respond to a voice lineup, which may affect that person's ability to identify the suspect. When you visited the police station, you listened to all 9 voices before making a decision. However, the police could have played one voice at a time and asked if you recognized the person or not. Researchers are currently investigating which type of procedure is best for voice identification.

Another system variable is called lineup composition. When choosing voices to be part of a voice lineup, it is important that the foil voices sound like the suspect's voice. This keeps the lineup fair because the suspect's voice does not stand out too much. However, including similar-sounding voices might make accurate voice identification quite tricky because it is difficult to distinguish between similar

voices. This could be why you found the voice lineup at the police station difficult.

The last system variable is recording length. The voices you heard in the voice lineup were each played for 1 min, which is quite a long time. You might imagine that listening to a longer recording would help you to compare each voice to your memory of the stranger's voice. However, our research has found that listening to shorter recordings does not increase the chances of being wrong [1]. In fact, it might even be *good* to include shorter voice samples because listeners do not get tired from paying attention to the voice recordings.

CONCLUSION

Do you think you would be a good earwitness? Based on everything you have now learned, do you think it is likely that you would be able to pick the stranger from the train out of the voice lineup? Think about how estimator variables and system variables could have affected your performance. Remembering a voice and identifying it from a lineup is not an easy task. There are lots of variables that might make voice identification difficult. It is very important that voice lineups are conducted in the best way possible, so that innocent people are not mistakenly identified as criminals.

ACKNOWLEDGMENTS

REFERENCES

1. Smith, H. M. J., Bird, K., Roeser, J., Robson, J., Braber, N., Wright, D., and Stacey, P. C. 2019. Voice parade procedures: optimising witness performance. *Memory*. 28:2–17. doi: 10.1080/09658211.2019.1673427
2. Johnson, E. L., and Helfrich, R. F. 2016. How brain cells make memories. *Front. Young Minds*. 4:5. doi: 10.3389/frym.2016.00005
3. Wells, G. L. 1978. Applied eyewitness-testimony research: System variables and estimator variables. *J. Pers. Soc. Psychol.* 36:1546–57. doi: 10.1037/0022-3514.36.12.1546
4. Yarmey, A. D., Yarmey, A. L., Yarmey, M. J., and Parliament, L. 2001. Commonsense beliefs and the identification of familiar voices. *Appl Cogn Psych.* 15:283–99. doi: 10.1002/acp.702
5. Cook, S., and Wilding, J. 1997. Earwitness testimony: Never mind the variety, hear the length. *Appl. Cogn Psych.* 11:95–111. doi: 10.1002/(SICI)1099-0720(199704)11:2<95::AID-ACP429>3.0.CO;2-O

6. Stevenage, S. V., Neil, G. J., Parsons, B., and Humphreys, A. 2018. A sound effect: exploration of the distinctiveness advantage in voice recognition. *Appl. Cognitive Psych.* 32:526–36. doi: 10.1002/acp.3424

YOUNG REVIEWER

ETHAN, AGE: 9
Ethan is a 9-years old boy that is currently in the fourth grade. He enjoys reading, writing, and studying math and history. During his free time, he plays basketball, tennis, and various other sports with his younger brother. He also loves exploring and learning about the nature and is currently a kid reporter. He lives in Virginia with his younger brother and parents.

AUTHORS

HARRIET M. J. SMITH
Dr. Harriet Smith is a psychologist at Nottingham Trent University. She studies how voice lineup procedures can be improved to make earwitnesses more accurate. She is also interested in how eyewitnesses recognize faces. Most of her work focuses on research, but she also teaches students. *harriet.smith02@ntu.ac.uk

NIKOLAS PAUTZ
Dr. Nikolas Pautz is a psychologist at Nottingham Trent University. He is interested in researching how humans are able to use different methods to learn and memorize new information. He is also interested in how people can remember older information and how their memories can be improved.

KATRIN MUELLER-JOHNSON

Dr. Katrin Mueller-Johnson is a psychologist and criminologist. She studies how psychological science can help the police interview witnesses better, so that witnesses remember more and find interviews less stressful. She also does research on how the police look after victims of crimes and what the police do to support those victims. She has taught police officers for 15 years at the University of Cambridge, and she now teaches criminology students at the University of Oxford.