



**UNIVERSITY
OF TURKU**

What is the required level of detailedness in asylum narratives for the case to be deemed sufficiently detailed?

Department of Psychology and Speech-Language Pathology, Faculty of Social Sciences

Master's thesis

Author:

Rebecca Heikola

15.3.2026

Turku

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin Originality Check service.

Master's thesis

Subject: Psychology

Author: Rebecca Heikola

Title: What is the required level of detailedness in asylum narratives for the case to be deemed sufficiently detailed?

Supervisors: PhD Jenny Skrifvars, doctoral researcher Rasmus Ahlgren and Professor Jukka Leppänen

Number of pages: 46 pages, 4 attachment pages

Date: 15.3.2026

Asylum interviews are used to evaluate if an asylum seeker faces persecution that meets the legal criteria for granting asylum. There are difficulties in assessing the asylum claims objectively. Guidelines direct asylum officials to use credibility indicators in their assessment, one of them being level of detail in the asylum seeker's narrative. Although high level of detail can be a sign of telling the truth, the problem lies with drawing the line in when someone is being detailed enough. In this study, the aim was to investigate if the number of details elicited from the asylum interview could predict the Finnish Immigration Service's assessment of detailedness. Secondary areas of interest were if the number of questions asked per topic, the proportion of open questions used and use of other credibility indicators affected the assessment of detailedness.

The sample of the study consisted of Finnish asylum decisions with their interviews from 2017-2018. The decisions were divided to topics, i.e. events or experiences that are part of the asylum claim. Each topic was coded with number of details, assessment of detailedness, question amount, proportion of open questions, use of other credibility indicators, case outcome and credibility of the topic.

A multinomial logistic regression model was created to test if the number of details could predict if the level of detail was deemed sufficient or not. The dependent variable was the Finnish Immigration Service's assessment of detailedness, the categories being: 1) sufficiently detailed 2) insufficiently detailed 3) detail is not mentioned. The predictors were number of details, number of questions asked, proportion of open questions and amount of other credibility indicators. There was no statistically significant connection between number of details and the assessment of detailedness. However, number of questions, proportion of open questions and amount of other credibility indicators did predict which category the topic belonged to.

The assessment of detailedness appeared to be subjective according to the results of this study, and the assessment was at chance level regardless of actual number of details elicited. This raised concern for the applicability of the criterion in credibility assessment. It should be used according to best practice standards, and its usability needs improving. The statistical analysis of this study had weaknesses, and all the results should be taken with caution and require replication, preferably with a new dataset. Still, this kind of research of the asylum system can help improve it towards a more objective and humane direction.

Key words: asylum seeker, asylum procedure, level of detail, credibility assessment, decision making, memory retrieval

TURUN YLIOPISTO
Psykologian ja logopedian laitos

HEIKOLA, REBECCA:
Pro gradu -tutkielma, 46 s., 4 liites.
Psykologia
Maaliskuu 2026

Turvapaikkapuhuttelua käytetään arvioimaan kohtaako turvapaikanhakija sellaista vainoa, joka vastaa turvapaikan saamisen laillisia vaatimuksia. On haastavaa arvioida turvapaikkahakemusta objektiivisesti. Suositukset ohjaavat virkailijoita käyttämään arviointiin uskottavuusindikaattoreita, joista tutkielman kohteena on yksityiskohtaisuuden taso turvapaikanhakijan kertomuksessa. Vaikka korkea yksityiskohtaisuuden taso voi olla merkki totuuden kertomisesta, ongelmallista on rajata, milloin joku on tarpeeksi yksityiskohtainen kertomuksessaan. Tässä tutkimuksessa tavoitteena oli tarkastella ennustaako turvapaikkapuhuttelussa esiintyneiden yksityiskohtien määrä Maahanmuuttoviraston arviota yksityiskohtaisuudesta. Toissijaisina kiinnostuksen kohteina olivat vaikuttavatko aiheesta kysytyjen kysymysten määrä, avointen kysymysten osuus ja muiden uskottavuusindikaattorien käyttö Maahanmuuttoviraston arvioon yksityiskohtaisuudesta.

Tutkimuksen otos koostui suomalaisista turvapaikkapäätöksistä puhutteluineen vuosilta 2017-2018. Päätökset jaoteltiin aiheisiin eli tapahtumiin tai kokemuksiin, jotka olivat osa turvapaikkahakemusta. Jokaiselle aiheelle koodattiin yksityiskohtien määrä, arvio yksityiskohtaisuudesta, kysymysten määrä, avointen kysymysten osuus, muiden uskottavuusindikaattorien käyttö, hakemuksen päätös ja aiheen uskottavuus.

Multinomiaalinen logistinen regressiomalli muodostettiin tarkastelemaan, ennustaako yksityiskohtien määrä sitä, oliko yksityiskohtaisuuden taso arvioitu riittäväksi vai ei. Riippuva muuttuja oli Maahanmuuttoviraston arvio yksityiskohtaisuudesta, luokkina: 1) riittävän yksityiskohtainen 2) ei riittävän yksityiskohtainen 3) yksityiskohtaisuutta ei mainittu. Selittäjät olivat yksityiskohtien määrä, kysymysten määrä, avointen kysymysten osuus ja muiden uskottavuusindikaattorien määrä. Tilastollisesti merkitsevää yhteyttä yksityiskohtien määrän ja arvion yksityiskohtaisuudesta välillä ei ollut. Sen sijaan kysymysten määrä, avointen kysymysten osuus ja muiden uskottavuusindikaattorien määrä ennustivat mihin luokkaan aihe kuului.

Tämän tutkimuksen tulosten perusteella arvio yksityiskohtaisuudesta vaikutti subjektiiviselta, ja arvio oli sattumanvarainen riippumatta puhuttelussa esiintyneiden yksityiskohtien määrästä. Tulos antaa syytä huoleen yksityiskohtaisuuden taso -kriteerin käytännön soveltumisesta uskottavuusarviointiin. Kriteeriä tulisi käyttää parhaiden käytäntöjen standardien mukaisesti, ja sen käytettävyys tarvitsee parantamista. Tämän tutkimuksen tilastanalyysissä oli heikkouksia, ja kaikki tulokset on ajateltava varovaisina arvioina, jotka tarvitsevat replikaatiota, mieluusti uudella otoksella. Silti tällainen tutkimus turvapaikkajärjestelmästä voi auttaa parantamaan sitä objektiivisempaan ja inhimillisempään suuntaan.

Asiasanat
turvapaikanhakijat, turvapaikkaprosessi, yksityiskohtaisuuden taso, uskottavuusarvio, päätöksenteko, mieleenpalautus

Table of contents

1	Introduction	5
1.1	Interviewing techniques	7
1.2	Alternative reasons for low level of details	9
1.3	Using level of detail for assessing credibility	11
2	Research questions	14
3	Methods	15
3.1	Materials	15
3.2	Procedure	15
3.2.1	Step 1: Case selection	15
3.2.2	Step 2: Coding of topics described in the decisions	18
3.2.2.1	Coding of topics in rejected cases	18
3.2.2.2	Coding of topics in granted cases	19
3.2.2.3	Exclusion criteria for topics	19
3.2.3	Step 3: Selecting relevant question-answer pairs in interviews	19
3.2.4	Step 4: Coding question type and number of details in answer	20
3.2.5	Step 5: Coding of other variables	21
3.3	Interrater analysis of the coding	22
3.4	Statistical analyses	23
3.5	Ethical concerns	24
4	Results	25
4.1	Descriptives	25
4.2	Multinomial logistic regression analysis	26
5	Discussion	29
5.1	Limitations and strengths of the study	33
5.2	Suggestions for further research	35
5.3	Conclusion	36
	References	39
	Appendix	43
	Appendix 1 Coding principles	43

1 Introduction

A refugee is a person outside of their home country who faces persecution there due to their ethnic background, religion, nationality, membership of a particular social group or political opinion, from which their home country cannot protect them from (United Nations, 1951). The UN Refugee Convention (1951) states that refugees have the right to freedom, non-discrimination, justice, work and education. An asylum seeker is a person who seeks international protection and can be granted refugee status based on their application, which means all refugees are asylum seekers before they are granted a refugee status (UNHCR, 2026). Currently, there are estimated to be 8.4 million asylum seekers globally (UNHCR, 2025). Evaluation of asylum claims is a complicated process, which, despite the guidelines that are in place, is not simple to conduct in an objective manner. In Finland the Finnish Immigration Service is responsible for evaluating asylum applications and determining whether the asylum seeker's life is in danger in their home country.

In Finland the asylum application is submitted to the police or border officials, typically upon arrival (Finnish Immigration Service, 2024). First, the Immigration Service gathers information about the applicant's past experiences and fears for the future, which is done in the asylum interviews (Finnish Immigration Service, 2024). Then the credibility of this information is assessed using the credibility criteria given by The European Union Agency for Asylum (EUAA). Based on the information elicited from the applicant, the Immigration Service evaluates if the applicant is in danger of future persecution in their home country (Finnish Immigration Service, 2024). Following the UN Refugee Convention, the legal criteria in Finnish law are that persecution must be due to ethnic background, religion, nationality, membership of a particular social group or political opinion, and that the applicant is not able or willing to ask for protection from their home country (Aliens Act, 2024).

There is rarely much physical evidence that can be reliably used to evaluate the claim, so the emphasis lies on the asylum interview and its credibility (EUAA, 2024; Kagan, 2003). Compared to other types of jurisdictions, such as the criminal law that requires certainty beyond a reasonable doubt, degree of certainty for asylum seeker's claims is allowed to be lower (Gyulai, 2013). To be granted asylum, the asylum officials only need to believe to a reasonable degree of likelihood that the asylum seeker would be harmed, if they returned to their home country (EUAA, 2024). Completely proving the need for asylum is not realistic, as the evidence is mostly gathered from the asylum interviews, and therefore only substantiating

the claim is required (Gyulai, 2013; EUAA, 2024). The credibility assessment is used as a tool to indicate whether the applicant is being truthful, and it aims to solve the problem of subjectivity and biases by applying a systematic evaluation towards the applicant's narrative (EUAA, 2024; Kagan, 2003). The asylum system is often thought to be abused by applicants with economic motives or false claims, which is why credibility assessments are in place, despite them containing plenty of room for subjective interpretation of guidelines as well (Kagan, 2003).

The EUAA has provided a recently updated guideline for asylum officials across the EU, including Finland (EUAA, 2024). The guide includes instructions about both how to assess the credibility of physical evidence and the oral statements of the applicant from the interview, as well as information about factors that can distort the oral evidence. As the result of the credibility assessment, each specific event or topic can be believed as factual or not (EUAA, 2024). Credibility is assessed separately for parts of the claim, such as different experiences, themes or events, and then in the final evaluation of claim, all these separate credibility findings contribute to the conclusion (EUAA, 2024). These events or topics can be such as "the applicant was repeatedly asked to join a terrorist group" or "the applicant's brother was killed by group x". Some events that are more relevant to the applicant's claim have more weight in the evaluation, and peripheral events or experiences should have less effect in the reasoning of the asylum decision (EUAA, 2024).

The EUAA (2024) presents five criteria for credibility assessments: internal consistency, sufficient level of detail, consistency with country-of-origin information (COI), external evidence and plausibility, i.e. realistic likelihood of something happening considering generally known facts. Out of the credibility assessment criteria, sufficient level of detail in answers is the focus of this study. The EUAA guide (2024) elaborates that true, personally lived experiences are more likely to be reported with a high number of details than experiences not personal to the interviewee, which is why level of detail is used in credibility assessments. This idea is supported by two meta-analyses on verbal indicators of truthfulness (Palena et al., 2021; Verschuere et al., 2021).

The information that is deemed credible is assessed with the legal criteria of asylum to conclude if the applicant will need asylum or not (EUAA, 2024). This means that asylum can only be granted based on the parts that the asylum officials believed to be truthful, even if other parts of the claim could also meet the legal criteria for asylum. Although the approach

may seem logical, it can create a risk for unfair treatment during different political climates or eras, when asylum officials collectively judge the applicant's credibility with higher standards than before (Vanto et al., 2022). In 2015 when Finland received a surge of asylum applications, many were rejected due to lack of credibility (Vanto et al., 2022). In comparison to previous years, Finnish Immigration Service officials assessed more applicants as not credible, and assessed similar claims inconsistently as well. This was possibly done to control the amount of people entering the country, as it was also politically incentivised (Vanto et al., 2022). The credibility assessment of asylum seekers is not only a societal and individual matter, but it is deeply rooted in psychological phenomena. Functioning of memory, especially that of emotional memories, individual and cultural differences in communication and lie detection are some of the key psychological aspects of a credibility assessment.

1.1 Interviewing techniques

As the whole asylum process depends heavily on the account of the applicant, it is important to use interviewing methods that are more likely to lead to more relevant and detailed information (Skrifvars, Antfolk et al., 2022). In an interview done according to best practice standards, the interviewer would ideally be listening more than speaking (Brandon et al., 2018). In the beginning of the interview, the applicant should be instructed to tell their story in as much detail as they can, and the interviewer's role is to listen (Brandon et al., 2018). More focused questions can be asked after the interviewee has finished their narrative. The interview should be done in an information-gathering style, focusing on seeking the truth (Meissner et al., 2012). Active listening and rapport building are the main methods in information-gathering style, for example by cues to continue speaking and showing respect and friendliness towards the interviewee (Brandon et al., 2018; Meissner et al., 2012).

These methods have been found effective to elicit information within the criminal context (Brandon et al., 2018). The information-gathering style itself also yielded better results in eliciting information even from reluctant interviewees (Vrij et al., 2014). It utilises open questions, which allows the interviewee to explain freely and only asks more questions after the interviewee has had a chance to answer completely. This gives the asylum applicant an opportunity for a long and detailed answers and encourages them to do so (van Veldhuizen et al., 2018). Though asylum interviews have different end goals than criminal investigations, both aim to elicit as much information as possible.

One of the key elements and methods in best practice interviewing is the focus on open questions – in the beginning of the interview a prompt for free recall should be used to investigate a topic of interest, such as someone’s reasons to seek asylum (Brandon et al., 2018; van Veldhuizen et al., 2018; Vrij et al., 2014). Open questions can be such as “Tell me why you seek asylum” or “Tell me about your family”, whereas closed questions are e.g., “What school did you go to?” or “Who threatened you?”. Using open questions can increase the amount key information and details from the interviewee (Snook et al., 2012), and the effect has support in the asylum context as well (Skrifvars, Antfolk et al., 2022). Cued invitations, that is, open questions that direct the interviewee toward a specific topic (e.g., “Tell me more about the car you mentioned earlier?”) are more informative than closed questions, when the applicant needs prompting to elaborate more (Skrifvars et al., 2020).

Mainly using closed questions is harmful, because the applicant should be given a chance to express detailed answers, if level of detail is used for assessing their credibility. According to Skrifvars and colleagues’ (2020) analysis of cases from 2017-2018 in Finland, a majority of the questions used in the interviews were closed questions, which can limit the amount of information in the answers. The questioning style can affect the detailedness of the whole interview, and therefore the credibility assessment of the asylum applicant. Depending on the type of questions posed and the number of additional questions, the answers might also stay on a general level without going into much detail, even if the applicant would be able to tell the narrative in a more detailed manner (Bodström, 2020).

The number of questions asked might be less critical than the questioning style itself, but it still affected the number of details when reporting experiences in Hope and colleagues’ study (2025). The number of questions could also be indirectly linked to how the Immigration Service assesses the applicant’s narrative, as a higher number of questions was present in cases where asylum was granted than in cases where asylum was rejected in Finland (Skrifvars et al., 2020). If the interviewer opted to cover every area of interest in the applicant’s claim and asked numerous questions, the interviewer’s approach to the case could lead to a high number of details and the applicant being assessed as credible. However, differences between interviewers’ approaches could then cause asylum decisions that depended on the applicant’s opportunity to provide a detailed narrative instead of their ability to do so.

1.2 Alternative reasons for low level of details

In addition to poor interviewing techniques, there can be numerous reasons why detail might be considered lacking by the asylum official. One of the most important factors affecting the number of details that the applicant can share is the functioning of autobiographical memory, the memory responsible for storing past events and experiences (Conway, 1996).

Autobiographical memories are interpretations of events that happened in a person's life, and they include not only sensory information but thoughts and feelings during the event (Conway, 1996). When autobiographical memories are retrieved from long-term memory, they are reconstructed rather than just retrieved, compared to a simple information unit such as a phone number (Conway, 1996). Memories can get altered when they are recalled, which means original details might change when the person talks about the memory (Cameron, 2010; Conway, 1996).

Especially when conducting an interview about events that happened years ago, the effect of forgetting and alterability of memories must be taken into account, therefore not expecting a perfectly detailed recount (Cameron, 2010). General ideas of memories are remembered in truthful retellings, but details are more prone to being forgotten. Particularly peripheral details are forgotten fast, and the number of details from memories decreases over time, then leading to memories becoming schematic representations of events that often happen instead of specific accounts (Conway, 2009; Herlihy et al., 2012). Peripheral details are also especially susceptible to changing over time, as well as aspects that are discussed with others (Cameron, 2010). This is necessary to remember in asylum interviews, since many memories might have been shared multiple times to multiple people before the asylum process, and often during the process as well, in repeated interviews (Herlihy & Turner, 2009).

High level of stress during an experience can also make it more difficult to remember peripheral details of it in an asylum interview (Gyulai, 2013; Herlihy et al., 2012). Memories of being persecuted can be very stressful or traumatic, and these are the main memories to be described in asylum interviews to establish why the applicant's home country would be unsafe. The asylum interview itself can be another source of stress or other negative emotions, and in a recent study by Skrifvars and colleagues (2025) most asylum seekers reported feeling nervous and afraid during the interview. In their study almost half of the asylum seekers reported that they were interrupted and felt like they needed to be concise, which is noteworthy regarding level of detail that could be elicited. They also felt uncomfortable

sharing personal experiences, especially ones they felt that they could be judged for, or the ones they could not remember everything from due to stress. Therefore, asylum officials need to understand why the level of detail might be low even in a truthful retelling of a distressing experience. Undetailed answers can be a sign of not remembering the event well or failing to retrieve it from memory (van Veldhuizen et al., 2018).

In addition to the functioning of memory, the behaviour of the applicant in the asylum interview can cause variance in detailedness. Even if the interview is conducted according to best practice standards, it cannot be assumed that asylum seekers are capable and willing to share many details of difficult experiences, such as trauma or possible taboos (Dowd et al., 2018; Gyulai, 2013; Herlihy et al., 2012). Topics that are not easy to discuss might vary from person to person, as well as in different cultures (Herlihy & Turner, 2009; Hope et al., 2022). Therefore, a completely detailed recount of everything that is relevant to the applicant's claim might be unattainable or arise later in the process, even if the applicant is truthful (Dowd et al., 2018). Simply the presence of an interpreter, which is often the case in Finland, was connected to less details in Ewens and colleagues' (2016) study, compared to speaking directly to the interviewer in their native language. Even though interpreters mostly translate accurately, any changes in the applicant's wording or style and especially mistranslations are a risk to the quality of the interview and the assessment of it (Keselman et al., 2010). However, the content or the situation itself is only one aspect that can affect the applicant's answering style.

Communication style and focus on self versus a social group are different between cultures, and this can affect the level of detail in interviews (Anakwah et al., 2020). In Western cultures, autobiographical memories tend to be more detailed, emotional and specific, revolving around the self instead of the social aspects (Herlihy et al., 2012; Wang, 2016). Closed questions that are heavily used in Finland do not give a lot of opportunities for implicit communication, which is common in collectivist cultures. This type of communication means less information (such as pronouns or tense) is expressed explicitly, but the message would be clear to a member of the same culture based on the context and implications in the message (Leal et al., 2018). In Leal and colleagues' (2018) study, this difference between explicit and implicit communication style affected the number of details when reporting experiences. They suggested that the explicit communication style requires the person to report more details for the interviewee to understand their experience. As the implicit communication style relies on context and less details are expressed, heavy use of closed questions increases the risk of

misunderstanding and limiting the amount of information that could be gained. There are also different ways to communicate emotion, and too little expressiveness can be a sign of not appearing credible to a Westerner, even though it could be the standard way of describing an emotional event in the applicant's culture (Hope et al., 2022). For better rapport-building and a higher quality interview, the asylum officials need to understand that the applicant's cultural norms and communication style can differ from theirs (Gyulai, 2013; Hope et al., 2022).

1.3 Using level of detail for assessing credibility

In normal life situations, others are expected to be speaking the truth. However, in an asylum interview the interviewer should evaluate the credibility of the claim and may thus be looking for signs of deception (Vrij et al., 2022). The signs should be verbal cues instead of nonverbal, as verbal signs are more accurate, one of them being high level of detail (Vrij et al., 2022). Two meta-analyses suggest that people telling the truth are more likely to include more details than people when they lie (Palena et al., 2021; Verschuere et al., 2021). This difference is also consistent across cultures, despite many other differences in deception cues (Vrij et al., 2022). Even if a lie is highly detailed, they tend to be less complex in nature to keep the cognitive effort of telling a false narrative lower (Leal et al., 2020).

However, professionals, including asylum officials, can only notice lying at a chance level (Bond & DePaulo, 2006). There is a common belief that lying could be detected from nonverbal cues, but in a meta-analysis exploring this, it appears people classify liars and truth-tellers correctly only 54% of the time, a percentage that is similar to simply guessing (Bond & DePaulo, 2006). Though detailedness, a verbal sign, can be a cue of truthfulness, there is a multitude of personal, situational and cultural factors affecting the level of detail in a single asylum interview. There are no clear cut-off points for when the interview is detailed enough or contains enough cues of truthfulness to be certain of the applicant's honesty (Vrij et al., 2022). In addition, the criterion of detailedness is difficult to use in practice during an interview, as the interviewer would need to keep the amount of detail in their mind while listening to the applicant, and try to maintain a good quality interview as well (Vrij et al., 2022). Another way to apply the criterion would be after the interview, counting the details from the interview transcripts, but this would be very time-consuming, and asylum officials are usually under time pressure to give out decisions (Vrij et al., 2022). Some of the detailedness also depends on the interviewer, as elaborated above.

Despite its potential in theory, studies suggest that the criterion of level of detail is used subjectively (Granhag et al., 2005; Maegherman et al., 2018). In Sweden, Migration board officials expressed not having a clear idea of how they make the judgement of someone's truthfulness, which highlights the subjective nature of credibility assessments (Granhag et al., 2005). In an experimental setting people seemed to use the credibility indicators together based on their general impression of the story and the applicant (Maegherman et al., 2018). This means people deemed the applicant not credible based on multiple credibility indicators when presented with a character and a story they did not consider deserving or generally trustworthy. Though Maegherman and colleagues' study (2018) did not use asylum officials as test subjects, it is possible that asylum officials are prone to similar use of heuristics and bias when assessing credibility.

In their work asylum officials cannot know if someone was being truthful or not, so there is no feedback on their decisions for officials to learn from and improve (Granhag et al., 2005). Experts in any field can make mistakes, especially if provided with context of the case that can cause bias, as researched on forensic fingerprint experts (Dror & Charlton, 2006). In an asylum interview, any facts related to the applicant or their story are contextual information that can make the official's assessment biased. The biases of the asylum official can then affect how they interview the applicant and how they assess the credibility, especially if the credibility is partly founded on the general impression of the applicant and their story.

Drawing the line for what is considered enough details is not simple, and asylum officials tend to expect the level of detail to be very high (Dowd et al., 2018; Skrifvars, Sui et al., 2022). What is considered as a significant or meaningful event is expected to be recounted without any detail being forgotten. However, the applicant's idea of a significant event can be different from the official's (Dowd et al., 2018), as well as the problems with relying on autobiographical memories in general, as described above. Similar results were found in a Finnish study by Skrifvars, Sui and colleagues (2022), where asylum officials expected the applicants to provide details such as times, names, locations and wordings in past conversations. An account not containing these descriptors was considered as a sign of untruthfulness by the officials. However, these are typically the kind of details that are forgotten easily and quickly after the event, if stored in the memory to begin with, which will lead people to give estimates or guesses when questioned (Cameron, 2010). For example, temporal information such as dates are usually not stored in memory, and people are typically able to give an estimate or a guess, but not the exact date. Yet, in credibility assessments for

asylum, this would be considered as a sign of untruthfulness despite being the normal way autobiographical memory functions. Because of this, Cameron (2010) recommended that the plausibility of the applicant remembering details, such as dates, should be considered in terms of importance: could the applicant be expected to purposefully remember the date of a certain event?

Although lack of detail cannot automatically be interpreted as a sign of deception, Skrifvars, Sui and colleagues (2022) found that Finnish asylum officials did not often consider alternative explanations for lack of details. Therefore, many possible reasons for a low level of details might not be explored in asylum interviews in Finland, meaning the credibility assessments would not be done according to best practice and the criteria of detailedness would be used without considering other explanations. Bodström (2020) suggests that concluding level of detail to be too low is used as a shortcut to conclude that the applicant was not credible, when the reason for low detailedness is instead in the questioning style: use of general questions and not asking for more elaboration.

These problems imply that the level of detail that is expected for an event to be assessed as credible is not realistic regarding the functioning of memory, as well as different cultural and personal communication styles. Finnish asylum officials seem to expect more detailedness than psychological research suggests is attainable, which leads to the risk of perceiving the applicant as untruthful when they are not. Therefore, research on how level of detail -criterion is used can be helpful in improving the quality of future asylum interviews.

2 Research questions

The objective of the current study was to investigate the level of detail required in an asylum seeker's account about specific events or experiences for it to be deemed credible. The secondary objective was to see if the proportion of open questions, number of questions, and the use of other credibility indicators affect whether the account was deemed credible based on level of detail. This study aimed to extend the studies by Skrifvars, Antfolk and colleagues (2022) and Skrifvars, Sui and colleagues (2022), by analysing cases of asylum decisions given in 2017-2018 from The Finnish Immigration Service.

If the number of details were connected to the Immigration Service's assessment of detailedness, it would imply that the criterion is used according to best practice guidelines and based on the details of the interview. This was my first hypothesis. However, I also expected there would likely be overlap in the number of details between sufficiently and insufficiently detailed topics due to the impracticality of the measure. Although a difference between assessment categories should appear in best practice, it can be difficult to use the criterion in a consistent manner. In asylum interviews, the interviewer would constantly need to keep the number of details in their mind, while still keeping the quality of the interview high (Vrij et al., 2022).

Secondly, I hypothesised that topics with a higher proportion of open questions would be more likely to be assessed as sufficiently detailed, due to open questions eliciting more details in previous research (Bodström, 2020; Skrifvars, Antfolk et al., 2022; Snook et al., 2012). This would imply that best practice interviewing techniques can help elicit more details, but it could be problematic for those assessed as insufficiently detailed, if questioning style caused the difference and limited the applicant's narrative. I assumed the number of questions asked about a certain topic might lead to more details about it, if the interviewer was approaching the topic from multiple angles and using open questions, but with closed questions it might just lead to repetition. Amount of other credibility indicators used for the topic was exploratively examined, because when used correctly, they should not have a connection to assessment of level of detail.

3 Methods

3.1 Materials

The data used in this study consisted of real asylum cases from the Finnish Immigration Service from 2017-2018. The cases consisted of the decision text and a complete interview transcript, in some cases transcripts of multiple interviews. The case files were anonymised and personal information was deleted by the Finnish Immigration Service. My study was an expansion for Skrifvars, Antfolk and colleagues' (2022) study and Skrifvars, Sui and colleagues' study (2022), which were based on a sample from the same cases obtained from the Finnish Immigration Service.

3.2 Procedure

3.2.1 Step 1: Case selection

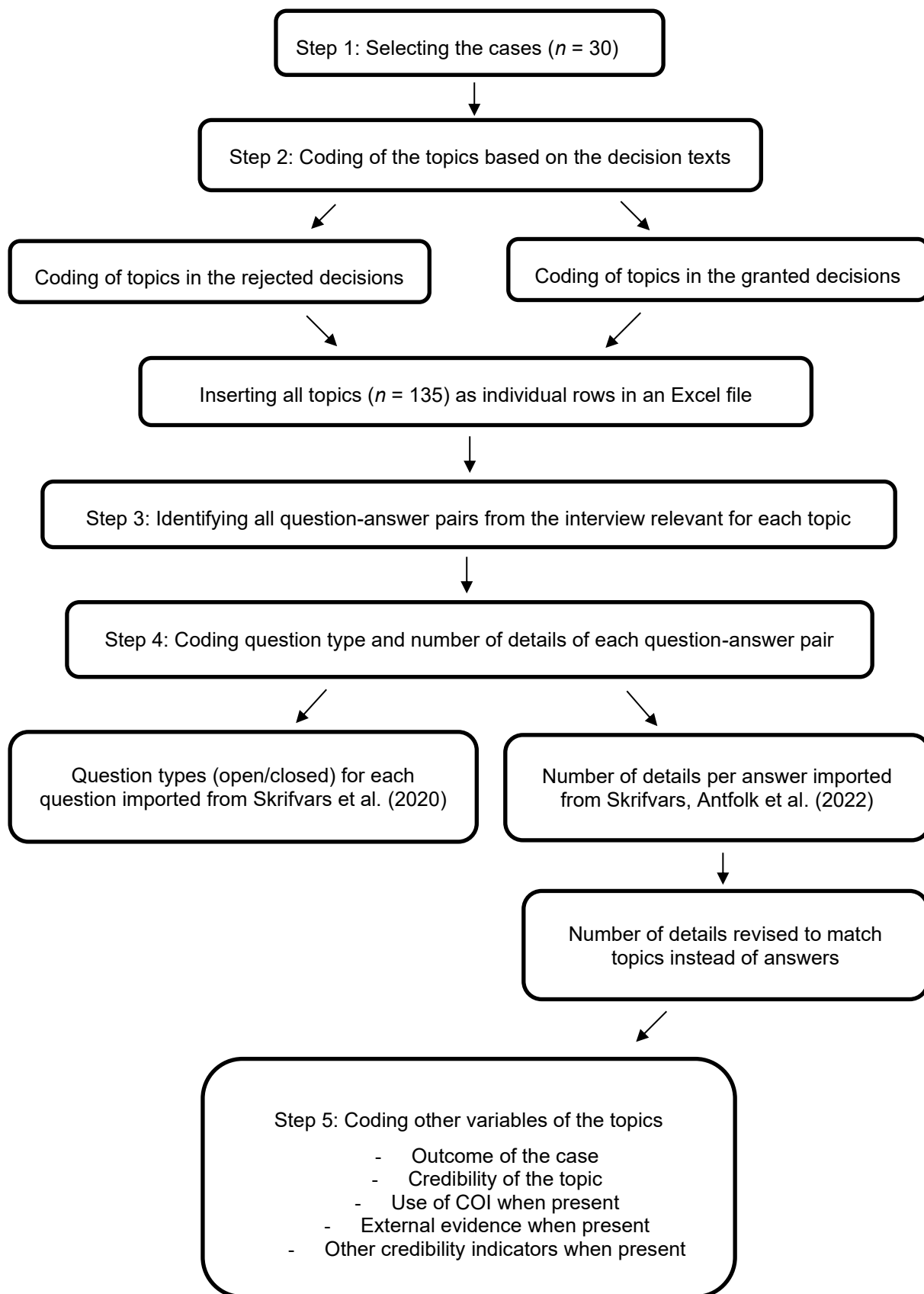
I selected the cases that had been previously coded in the publications Skrifvars, Antfolk et al. (2022) and Skrifvars, Sui et al. (2022) to be included in the current study. I included cases where asylum was rejected ($n = 35$) from both studies. In Skrifvars, Sui et al. (2022) the cases were coded for the use of credibility indicators, and in Skrifvars, Antfolk et al. (2022) they were coded for the number of details. All cases where asylum was granted ($n = 29$) from the study Skrifvars, Antfolk and colleagues (2022) were added to the sample of this study. The sample of Skrifvars, Sui et al. (2022) had not included any cases where asylum was granted.

Based on the starting sample ($n = 64$) I decided to exclude some cases that for different reasons were not suitable for the planned analysis. The decisions which had first been rejected and later accepted only due to a court appeal without a second interview, were excluded from analysis ($n = 11$). This is because they cannot be counted to be factually accepted by the Finnish Immigration Service, as their original negative decision was simply overruled by the court. They could not have been coded as rejected either, because they did not contain the information of reasons for rejection or the credibility indicators used, so they were not comparable to other rejected cases. However, applications where the Finnish Immigration Service accepted that the applicant would be in danger in their home country but suggested internal fleeing, then appealed in court and granted asylum due to court decision that internal fleeing was not possible, were included in the analysis. This is because the danger itself was

believed by the Finnish Immigration Service, which means it at least some of the topics were accepted as credible.

Cases whose claim was either religious conversion or belonging to a sexual minority, were excluded ($n = 10$). The type of questions to ask about identity are different than questions regarding specific events in cases with political persecution as the claim. For example, in asylum applications in Finland where the claim was based on sexual orientation, the interview questions were focused on past relationships, feelings and development of identity instead of the risk of persecution (Selim et al., 2025). In my data this difference was prominent upon attempting to code the identity-based cases objectively by topic. They had broader topics that were thoroughly investigated by multiple questions, often in multiple interviews, and the topics were heavily overlapped. This would have made the data unbalanced and the coding more subjective, which is why I chose to exclude them.

Cases where credibility was not the issue, were excluded, because they are not relevant for the research question ($n = 5$). These are cases that the Immigration Service explicitly stated to be groundless based on a safe country of origin or exploiting the asylum system. Any topics or sometimes complete cases that were shared between family members or couples were excluded, if I did not have everyone's interviews, because it made analysis of complete topics impossible ($n = 8$). Shared topics could be for example events that happened with multiple applicants present, such as an attack to the family house, when other family members are referred to in the decision text. Then the Finnish Immigration Service would combine information from all the applicants to come to their conclusion on credibility. The final number of cases included in the analysis was 30, consisting of 18 rejected and 12 granted cases, two of which received secondary protection status. Coding after case selection is summarised for clarification in Figure 1. My coding principles are in Appendix 1.

Figure 1*Flowchart of the coding process*

3.2.2 Step 2: Coding of topics described in the decisions

In the next step, I read through the included decisions and identified the topics of the application. In this study, I defined topic as an event or fact that was part of the asylum seeker's claim for seeking asylum that the Finnish Immigration Service mentions in the decision. I approached the research questions through these topics, because topics are the areas of the applicant's claim that the Immigration Service used credibility indicators for and named the indicators used (such as detailedness). All topics were coded as individual rows in an Excel document as a short description of the topic, e.g., "The Immigration Service does not accept that you received a threat call" (rejected case) or "You were kidnapped by a militant organisation" (granted case). Then each topic was assigned to one of the three categories based on the Finnish Immigration Service's conclusion of detailedness for that topic: 1) the level of detail was sufficient 2) the level of detail was insufficient 3) the level of detail was not mentioned for the topic.

The content and structure of the Finnish asylum decisions depend on whether the applicant was granted asylum or rejected. In granted cases, the officials were not required to explain their reasonings for granting asylum or give any evaluation of the credibility of the applicant's claims. Therefore, granted cases only briefly listed the topics of an applicant's claim, without mentioning detailedness or any other credibility indicators. The coding of topics was done based on the information mentioned in the decisions, not the content of the interviews.

3.2.2.1 Coding of topics in rejected cases

The coding of topics in rejected cases was directly imported from the coding in the study by Skrifvars, Sui et al. (2022). To make analysis of rejected and granted cases more equal and the variation within them similar, I included all topics from rejected cases, even when detailedness was not mentioned. A topic was based on any phrase where the Finnish Immigration Service declared their opinion on something the applicant had stated in their interview. These phrases were such as: "the Finnish Immigration Service accepts/does not accept..." and "the Finnish Immigration Service finds...", where the name of the organisation was present.

3.2.2.2 Coding of topics in granted cases

As the granted cases did not contain any reasonings for the outcome, it was not possible to determine with certainty which topics mentioned in the decision the official deemed to be credible and accepted as factual. Because of this, I decided to only include topics that were related to the stated legal ground for granting asylum to the applicant (e.g., political opinion or ethnicity). If the decision included additional topics related to other grounds for asylum (e.g., religion), these were not included in the analyses.

One full sentence describing reasons to apply for asylum counted as one topic, except when one sentence was a list of multiple unrelated events because of the Finnish Immigration Service's concise wording. As only the main parts of the application were recounted in the decision, the coding was different from rejected decisions, and there was more variation in the amount of information presented in the decision text in granted cases.

3.2.2.3 Exclusion criteria for topics

Each topic was coded only once, even if it was mentioned in the decision multiple times, such as a part of a conclusion. The criteria to exclude a topic were 1) the topic was a shared experience between family members and not all interviews were available 2) the topic being accepted/not accepted as factual was the person's nationality and background, because it is not assessed in the asylum interviews but earlier in the process 3) the topic was disregarded based on previous topics that were also not accepted, not by what the applicant had said 4) the topic was not investigated in the asylum interview despite being mentioned in the decision text 5) if a topic was of similar theme to another topic and the questions were heavily overlapped, these two were combined to a single topic. With these criteria, the total number of topics to analyse was 135.

3.2.3 Step 3: Selecting relevant question-answer pairs in interviews

After establishing the topics from the rejected and granted cases, I identified the question-answer pairs about the topics in question. The inclusion criteria for question-answer pairs to be included in a topic were: 1) the question and/or answer mentioned the topic (regardless of it containing any new information or not) or 2) the topic was not directly mentioned but implied by the context (such as previous or following questions clarifying which topic it was). Exclusion criteria for question-answer pairs were: 1) the topic was mentioned only as a time

indicator for another event (e.g., applicant stated they got into an accident one day after their mother's death) or 2) it was unclear by the context which topic was discussed in the question-answer pair and no new information was presented.

3.2.4 Step 4: Coding question type and number of details in answer

For the included question-answer pairs, the question types and number of details in answers were mainly retrieved from the previous coding for studies Skrifvars and colleagues, 2020 and Skrifvars, Antfolk and colleagues, 2022. Classification of question type was imported directly, and it included 10 options for the type. These were: invitation, cued invitation, focused question, forced choice, yes/no -question, suggestive question, other/unclear, utterance, summary and type not applicable. However, for this study I combined invitation and cued invitation to be coded as open questions, which was the main area of the interest. Open questions in these interviews are such as: "Tell me why you are seeking for asylum" or "Tell me about the shooting". Then I calculated the percentage of open questions asked within the topic as a new variable.

Number of details was counted as information units from basic statements, a coding scheme originally developed by Allwood and colleagues (2005). This coding scheme was used as a base for operationalising level of detail in the study by Skrifvars, Antfolk and colleagues (2022), which I imported the number of details from. A unit of information in the asylum interviews was a basic statement about the events, thoughts, individuals, or anything else relating to the topic. A sentence could include multiple basic statements or descriptions of the topic. These could be such as "I went to my mother's house" as one unit or "I went to my mother's house with my sister" as two units. The information was counted as a unit only the first time it was mentioned. In the study by Skrifvars, Antfolk and colleagues (2022) the information units had been calculated for each question-answer pair. For the purposes of this study, I combined the number of information units from several question-answer pairs to correspond to a topic instead. However, this could not be done for answers where the applicant talked about more than one topic. The total number of information units for answers that included information about several topics were revised, so that the number of information units corresponded to the individual topics, rather than to the full answer. Some details overlapped between topics, such as the same group of people threatening the applicant in different situations or interconnected events. Details that corresponded to several topics were included in all those topics.

3.2.5 Step 5: Coding of other variables

Other variables that I coded from the decision texts were outcome of the case (rejected/granted), credibility of the topic (credible/not credible/not available) and other credibility indicators than detailedness (several used/one used/none). These were coded for each topic. Apart from case outcome, all other variables were only present in rejected cases. The credibility of the topic refers to whether the Immigration Service concluded it to be credible or not, based on all the credibility indicators they used in evaluation of the topic. The credibility of the topic was coded as accepted as credible or not accepted as credible. Credibility was not mentioned in granted decisions, so for those it was coded as not available. Other credibility indicators the Immigration Service mentioned in the decision regarding a topic's credibility were coded as their amount per topic (none, one, several). These indicators were such as consistency, personality of threat or durability of threat.

Use of COI (supports the applicant/against the applicant/unclear/not used), external physical evidence (present/not present) were also coded for each topic, but they were not included in the analysis of this study. COI was mentioned in assessment of only 22 topics, of which in 7 COI was in favour of the applicant, in 11 against the applicant and in 4 the implications of COI were unclear. There was no significant connection to the assessment of detailedness, $p = .091$, Fisher's exact test, indicating that COI was used similarly regardless of whether the topic was deemed sufficiently detailed, insufficiently detailed or when detailed was not mentioned. COI was not included in further statistical analyses because of this. External physical evidence was also present in 22 topics. External evidence was similarly discarded from further analysis, as it was not statistically connected to the assessment of detailedness, Pearson $\chi^2(2) = 4.74$, $p = .094$. Like COI, according to the test external evidence was used regardless of how the Immigration Service had assessed the level of detail for the topic. All coded variables are elaborated in Table 1.

Table 1*Coded variables*

Name of variable	Description	Measurement
Case outcome	Whether the applicant was granted asylum or the claim was rejected	1 = asylum granted, 2 = asylum rejected
Credibility of the topic	The Immigration Service's conclusion of the credibility of the topic	1 = credible, 2 = not credible, 3 = not available
Assessment of detailedness	The Immigration Service's assessment of the level of detail of a topic	1 = sufficient, 2 = insufficient, 3 = not available
Other credibility indicators	The number of other credibility indicators used for credibility assessment (not including detailedness)	0 = zero, 1 = one, 2 = several
COI	The use of COI (explicitly mentioned) in credibility assessment	1 = in favour of applicant's narrative, 2 = against applicant's narrative, 3 = unclear, 4 = not available
External evidence	The use of external physical evidence in credibility assessment	1 = yes, 2 = no
Open questions	The proportion of open questions in relation to all questions about the topic	Numeric (percentage)
Number of questions	The number of questions asked about the topic	Numeric
Number of details	The number of information units elicited from the topic	Numeric

3.3 Interrater analysis of the coding

An interrater analysis was conducted to investigate the coding principles' clarity and usability. The other rater coded a random sample of 19 cases from the materials of the study. The other rater's coding was based on the same coding principles that I used. I analysed our degree of agreement for the question-answer pairs from each topic with Cohen's kappa. They are the focus of the interrater analysis, because the question-answer pairs are what determines the number of details in the topic. The question-answer pair numbers were coded as data points for both raters, and when one of the raters did not add the same question-answer pair, it was

coded as 0 for that rater. I also calculated Cohen's kappa in a similar manner for the topics to check the agreement between raters for their coding.

There was a substantial agreement between me and the other rater for question-answer pairs, $\kappa = .65$, 95% CI [0.62, 0.69], $p < .001$. The degree of agreement for topics was substantial as well, $\kappa = .77$, 95% CI [0.69, 0.86], $p < .001$. This means we agreed above chance level, i.e. the agreement did not simply happen due to chance. Although the coding was subjective, a substantial agreement means the coding principles work to a certain extent, allowing some reliability even within different raters. The 95% confidence interval for question-answer pair agreement was narrow due to the large sample ($n = 1004$) of question-answer pairs in the analysis. A small confidence interval implies that the statistical value is quite robust, and the agreement is indeed substantial, even at the lower end of the confidence interval.

3.4 Statistical analyses

In statistical analyses, the aim was to investigate how many details are required for different assessments of credibility based on detailedness: 1) level of detail is sufficient 2) level of detail is insufficient 3) level of detail is not mentioned. To test these differences, I conducted a multinomial logistic regression analysis. First, I explored if the assumptions for multinomial logistic regression were met. The data for number of details and number of questions asked were skewed, but I decided to keep the data in its original form, as the outliers are a natural part of this data to be included. To test for multicollinearity between the predictors, I checked correlations between them. Correlations were small (below .20), except number of questions and number of details had a moderately high correlation, $r = .68$, $p < .001$. This indicated that when interpreting the results, the unique effects of number of questions and number of details can be difficult to separate, and warrants caution. The assumption about linear connections of the predictors to the logit of the dependent variable was also fulfilled, except for the variable number of details ($p = .033$). Again, this indicates that the results should be interpreted with caution. One of the categories, sufficiently detailed, consisted only of 9 instances. Despite the small size of this category, I decided to use multinomial logistic regression, because there is no suitable non-parametric test to conduct this analysis, and the other assumptions were mostly fulfilled. However, interpreting the results from this analysis must be done with caution, since the assumptions were not completely fulfilled. The results derived from a small sample category might not be representative of the population. Overall, all the aforementioned

limits indicate that the results of the analysis should be interpreted carefully, and mainly considered as possible indications that need future replication.

I created a multinomial logistic regression model to see which variables would predict belonging to these three categories of different levels of detail. The goal of the regression analysis was to predict belonging to sufficient detail or detail not being mentioned categories compared to the category where detail was insufficient. The dependent variable in the regression model was the assessment of detailedness (sufficient, insufficient or not mentioned). The main predictor was number of details, and the covariates were number of questions asked about the topic, the proportion of open questions and amount of other credibility indicators used. Case outcome (asylum accepted or rejected) was included in the model as a factor. The aim was to see whether the number of details elicited from the answers could predict if the level of detail was deemed sufficient or not. To answer the other research question of whether question type and question amount affect the assessment of detailedness, proportion of open questions and number of questions asked about the topic were included in the model as covariates. Amount of other credibility indicators used for assessment of the topic was also added to see if it would affect the assessment of detailedness. All analyses were conducted on IBM SPSS 30.

3.5 Ethical concerns

The materials have been obtained from the Finnish Immigration Service by request. Personal information was deleted by the organization before classifying the documents publicly available. Ethical permission for collecting the data has been received from Research Ethics Board for Psychology and Speech and Language Pathology of Åbo Akademi University.

4 Results

4.1 Descriptives

This study included 135 topics from 30 cases. The minimum number of topics in a case was 1 and the maximum number of topics was 8. Of all the topics, 9 were assessed as sufficiently detailed, 41 as insufficiently detailed and in 85 topics, detail was not mentioned. The average number of details for each category is presented in Table 2. The distributions of number of details per category are visualised in Figure 2, each dot representing a topic. A total of 22 topics were deemed credible, 66 were found not credible and for 47 topics credibility was not explicitly mentioned due to the case being granted asylum. Credibility of the topic and case outcome were statistically connected to each other, $\chi^2(2) = 104.27, p < .001$. Average number of details separated by case outcome is also presented in Table 2. In cases where asylum was rejected, the topics were mostly not accepted to be credible either. In the two cases granted with secondary protection, 3 topics were deemed credible, but 4 were not. Other granted cases did not contain information of credibility. All 9 topics that were concluded to be sufficiently detailed, were also deemed credible, and all the insufficiently detailed topics were deemed not credible, meaning the credibility of the topic matched the assessment of detailedness.

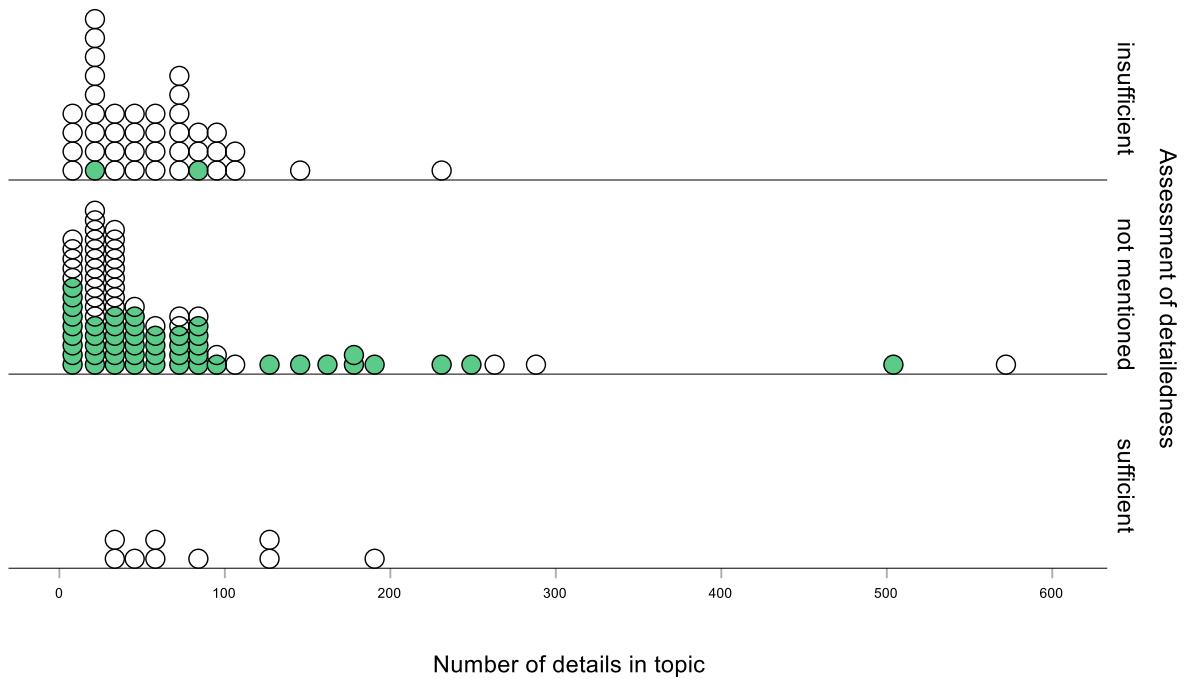
Table 2

Average number of details separated by assessment and case outcome

Assessment/Decision	Average number of details	Median	SD	Min	Max
Sufficiently detailed	87.78	55	53.54	32	189
Insufficiently detailed	57.88	50	42.28	8	226
Detail not mentioned	69.46	37	95.19	3	572
Asylum granted	72.43	47	84.14	3	504
Asylum rejected	63.32	38	77.87	8	572

Figure 2

Distribution of number of details by assessment category



Note. Topics where asylum was granted are depicted in green.

Number of questions ranged from 1 to 152, and the average number of questions asked about a topic was 14.14 ($SD = 17.32$). The average percentage of open questions was 22.54% ($SD = 18.96\%$). Amount of other credibility indicators was a categorical variable with value options for none = 0, one = 1, several = 2. Almost half, 49.63% of all topics did not use any other credibility indicators. These consist of both topics that were assessed only with detailedness and topics from granted cases that had zero credibility indicators mentioned. One credibility indicator (or one and detailedness) was used for 39.26% of topics, and several indicators were used for 11.11% of topics.

4.2 Multinomial logistic regression analysis

The aim of the study was to see if there was a certain number of details connected to the assessment of credibility based on detail, along with investigating if proportion of open questions, number of questions and use of other credibility indicators affected the assessment of detailedness. These questions were approached by creating a multinomial logistic regression model.

The regression model was better at explaining belonging to categories than an empty model, $\chi^2(10) = 72.89, p < .001$. However, the predicted values differed significantly from actual values, Pearson $\chi^2(258) = 300.78, p = .035$. The model sorted correctly 74.1% of instances, and the explanatory power was moderate, Nagelkerke $R^2 = .51$. This indicated that the model was partially effective, in that it was significantly better than no model, and explained a decent amount of variation, but did not perfectly fit the data.

According to the likelihood ratio test, the number of details within a topic did not significantly predict the assessment of detailedness, $\chi^2(2) = 0.39, p = .824$. Neither did the percentage of open questions, $\chi^2(2) = 5.15, p = .076$. However, the number of questions asked about the topic, $\chi^2(2) = 8.37, p = .015$, and amount of other credibility indicators used, $\chi^2(2) = 7.86, p = .020$, did significantly predict the assessment of detailedness. The outcome of the case was also a significant predictor for the assessment, $\chi^2(2) = 64.83, p < .001$.

The main predictor, number of details did not increase or reduce odds of the topic being assessed as sufficiently detailed, or detail not being mentioned in credibility assessment compared to topics that were assessed as insufficiently detailed. A larger number of questions asked about the topic increased the odds for it to be assessed as sufficiently detailed rather than insufficiently detailed. Number of questions asked did not increase or reduce odds to belong in detail not mentioned category when compared to insufficiently detailed category. More open questions increased the odds of being assessed as sufficiently detailed, but not detail not being mentioned compared to insufficiently detailed category. More credibility indicators being used increased the odds of detail not being mentioned compared to being assessed as insufficiently detailed. Statistics of all variables are reported in Table 3.

As a sensitivity analysis, I created an additional regression model without number of questions as a predictor to test if the effect of number of details would change, as number of questions and details had a correlation. Similarly to the main regression model, number of details did not increase or reduce odds for the topic to be assessed as sufficiently detailed, $OR = 1.01, 95\% CI = [0.99, 1.01]$, or the detail not being mentioned, $OR = 1.00, 95\% CI = [0.99, 1.01]$, when compared to insufficiently detailed topics.

Table 3*Statistics of multinomial logistic regression analysis*

Variable	Sufficient detail vs. insufficient detail					Detail not mentioned vs. insufficient detail				
	<i>B</i>	<i>SE</i>	<i>p</i>	<i>OR</i>	<i>95% CI</i>	<i>B</i>	<i>SE</i>	<i>p</i>	<i>OR</i>	<i>95% CI</i>
Number of details	-0.00	0.01	.787	1.00	[0.99, 1.01]	0.00	0.00	.721	1.00	[0.99, 1.01]
Number of questions	0.10	0.04	.013*	1.10	[1.02, 1.19]	-0.02	0.02	.437	0.98	[0.95, 1.03]
Amount of other credibility indicators	0.52	0.70	.459	1.68	[0.42, 6.69]	1.22	0.47	.010*	3.38	[1.34, 8.49]
Percentage of open questions	4.73	2.07	.022*	113.09	[1.97, 6488.88]	1.25	1.34	.352	3.47	[0.25, 47.77]

*Statistically significant

5 Discussion

This study aimed to investigate whether number of details in asylum interviews was connected to the Finnish Immigration Service's assessment of credibility based on level of detail. Other associations to the assessment of detailedness that I focused on were proportion of open questions, number of questions and use of other credibility indicators for assessment. According to the statistical analysis, the proportion of open questions, the number of questions and use of other credibility indicators predicted how the Finnish Immigration Service had assessed the applicant's credibility with the criteria of detailedness. However, the results should be interpreted with caution due to the limitations of the statistical tests.

Firstly, I wanted to see whether the number of details in the applicant's interview was connected to the Immigration Service's assessment of detailedness. According to the regression model the actual details and the assessment of it did not have a statistical connection. This result did not give support for the hypothesis that the level of detail is connected to the assessment of detailedness, thereby challenging the theoretical ideal of how credibility assessment should work. Regardless of how many details the applicant provided in their account, the Immigration Service's decision of whether it was sufficient or not was at chance level. Since there were no statistically significant differences between the assessment categories, there were no observable cut off points for something to be detailed enough. Assessing detail was subjective in nature, as Vrij and colleagues (2022) also suggest in their research. They also highlight the difficulty of using the criterion of level of detail in practice. This could explain why there was no statistical connection to actual number of details, when the criterion is simply hard to use accurately. The assessment of detailedness was not based on the calculated number of details elicited, and it would be very time-consuming for the Immigration Service to conduct such calculations for a single credibility indicator for every applicant.

Other previous research implies that the level of detail -criterion is used subjectively too (Bodström, 2020; Granhag et al., 2005; Maegherman et al., 2018), and the results of this study support that notion. Similarly to Bodström's (2020) analysis of Finnish asylum interviews, it appears that the criterion of level of detail was not used consistently, and it was not based on the actual number of details. Bodström suggests that the Finnish Immigration Service might state the applicant is not credible based on low level of detail and then use it as a means to reject the application. This would explain why level of detail was often mentioned in the

asylum decisions, but it did not have statistical connection to how many details the applicant's answers included. Several cases in this study were asylum applications filed in 2015-2016 (interviews from 2017-2018), which was when Finland was receiving higher numbers of asylum seekers than before. As Vanto and colleagues (2022) theorise, collectively shifting to a stricter assessment of credibility indicators could be used as a method to keep the numbers of refugees in Finland lower, which was also the government's publicly stated goal at the time.

Another issue that could have affected the assessment of level of detail is the high expectations for how detailed the narrative should be (Dowd et al., 2018; Skrifvars, Sui et al., 2022), and it could have led to assessing more topics as insufficiently detailed than what the number of details indicated. Asylum officials may also not know how much details are usually forgotten over time and therefore expect more. Furthermore, all asylum officials have their own idea of what constitutes as enough details, which could have caused some topics with a similar number of details to be assessed differently.

The second research question focused on whether the question type, question number and use of other credibility indicators affect the assessment of detailedness. As hypothesised, a higher proportion of open questions statistically predicted being assessed as sufficiently rather than insufficiently detailed – this would imply that open questions indeed help elicit more information. However, the 95% confidence interval for the value was very large, so the result should not be taken as a robust finding. Still, use of open questions has strong support in best practice interview techniques and has been shown to be effective both in criminal investigations and in the asylum context in previous studies (Brandon et al., 2018; Meissner et al., 2012; Skrifvars, Antfolk et al., 2022; Snook et al., 2012; van Veldhuizen et al., 2018). My results regarding open questions lend support to this idea and the use of open questions in asylum interviews, particularly since the interview content is hoped to be as detailed as possible.

In general, open questions were not used much in the sample of this study. Still, they were always the starting question of the asylum interview, which is in line with best practice of information-gathering interviewing styles (Brandon et al., 2018; Meissner et al., 2012). The cases did not have a lot of variation in the proportion of open questions, as the interview structure was similar for all cases. However, some topics might have had a greater proportion if they were mostly described in response to the first question ("Tell me why you are seeking

asylum.”). If the applicant explained a topic in a few answers instead of spreading it out in several closed questions after the first open one, the proportion of open questions per topic was naturally larger. The applicant’s answering style could also be dependent on what they are willing to share (Dowd et al., 2018; Gyulai, 2013; Herlihy et al., 2012) or the typical communication style in their culture, such as favouring implicit communication (Leal et al., 2018) or focusing on the group level (Anakwah et al., 2020; Wang, 2016). Closed questions allow less implicit communication than open ones (Leal et al., 2018), so mostly asking closed questions could then reduce the opportunity to provide details even more.

Interviewers have their own differences as well, and some might be more trained or motivated to follow best practice techniques and utilise open questions to gain a full narrative before specifying details by directing the applicant with closed questions. Since closed questions were heavily used in these interviews, it could cause the effect of open questions to be small in this data. On a topic-level mostly using closed questions could have led to the applicant providing less details, which could explain why the topics with less open questions were more likely to get assessed as insufficiently detailed. To summarise, both the interviewing style and the applicant’s personal style of answering and explaining their claim affected the proportion of open questions.

In the regression model, number of questions asked predicted being assessed as sufficiently detailed compared to insufficiently detailed. The odds ratio was small, so the difference one question made for the assessment of detailedness was not large either. This appeared logical, especially regarding answers to closed questions, which might not have contained many new information units. Number of questions was still a significant predictor for the regression model. There are many possible reasons for this connection. It could be that the number of questions hid the effect of number of details on the assessment, as number of questions and details correlated substantially. If more details were elicited when more questions were asked, it could have increased the odds to get assessed as sufficiently detailed. Still, number of details was not a significant predictor in the additional regression model with the variable question number removed. As number of details on its own had no statistical connection to the assessment, simply asking more questions was what made the difference according to the regression models. For example, the asylum official might have seen that a topic had been investigated thoroughly with multiple questions, and with that reasoning be satisfied with the level of detail in the topic.

The range of number of questions per topic was large, which was partly due to some of the topics being less important for the asylum claim than the others. For example, a topic such as a family member working with a Western country could simply be mentioned in one question and answer. However, central topics related to persecution were investigated thoroughly, sometimes in multiple interviews as well. Some topics consisted of broader statements, such as the applicant being the target of an extremist group for a long time, and some were more focused on a specific event, e.g. an individual threat letter. This large variance in number of questions is necessary to keep in mind when interpreting the results of the regression analysis. The small number of questions asked within a topic could have been open questions, such as the starting question of the interview. Therefore, a small number of questions combined with a high number of details might not reflect the assessment of detailedness. Some of the answers to the starting question were very long, and then most of the information units came from the answer to that question. However, that does not account for assessing the topic as insufficiently detailed, if more elaboration was not asked for. The topics with fewer questions being assessed as insufficiently detailed would indicate that the applicant was not given the opportunity to explain in a detailed manner (if the questions were mostly closed), or they were not asked to specify after the open questions.

It seems problematic that the number of questions about the topic predicted whether it was deemed detailed enough. If officials focused their questions only on the topics which they found interesting and relevant, there is a risk for bias towards the topics the official had chosen. These topics were then the ones that they deemed credible enough based on detail. Yet, the actual number of details in the topic did not affect this, but the connection was instead to the number of questions the official chose to ask about it. It is understandable that the official can focus on certain areas of interest and topics related to the asylum claim more than peripheral aspects of the applicant's story, but then the credibility indicator in use should not be level of detail. Important topics could be accepted to be credible based on consistency or threat being very personal or acute. This also implies that level of detail was used subjectively in credibility assessment, with no straightforward connection to actual detail.

Amount of other credibility indicators used in assessment was also a significant predictor in the regression model. However, as it only predicted belonging to the category where detail was not mentioned, it simply seems that other credibility indicators were used, if detail was not. This is how credibility indicators were intended to be used, so these results were not surprising.

5.1 Limitations and strengths of the study

A notable limitation in the statistical analysis of this study was the small category size of sufficiently detailed topics. Not all requirements of multinomial logistic regression were fulfilled, but it was used nevertheless. This weakened the statistical strength of the regression model, and all results and their interpretations need to be understood as possible implications rather than strong support for any theory or previous research. It is possible that number of details did not observably predict the assessment of detailedness because of the small sample size, and the effect remained unseen in this dataset. Although quantitative research has its perks, fitting this kind of source material into a statistical format was somewhat artificial and led to simplifying complicated and multileveled material. Yet, statistical results gained from this study can bring valuable information to create ground for improving the asylum system and shed light on new areas to research. Additionally, the substantial agreement in the interrater analysis regarding the questions chosen for each topic supports the validity of the coding principles for the number of details.

The methodology itself came with challenges caused by the nature of the materials. Cases where asylum was granted and where it was rejected had greatly different decision texts and descriptions of the asylum claims. Only rejected cases and one case with secondary protection were thoroughly explained and evaluated with credibility indicators that were mentioned. There was plenty of variation within what was written in the decisions, which I tried to code as consistently as possible to become topics for the study. This variation was mostly due to the difference between rejected and granted cases, but also within similar outcomes as asylum officials had their own styles for wording the decisions. However, the substantial agreement in the interrater analysis supports that the coding principles for topics were clear enough. The difficulty in consistent and objective coding appeared especially in identity-based claims, which I chose to exclude from the study. This decision reduced the sample size and the applicability of the results, as the sample consists only of claims based on political persecution.

In this study I was interested in the Finnish Immigration Service's point of view and how it deems an applicant to be credible with the level of detail that was present in the interview. For granted cases, any credibility criteria, including detail, was never mentioned in the decision. Therefore, I cannot compare the assessment of detailedness between granted and rejected cases. I can only know the general level of detail that is present in those applications that were

granted asylum. In addition, the granted cases were highly varied in their number of details, as highlighted by the range of the detail not mentioned category. This category also included topics from rejected cases, which had similar variance and were added to the category to balance it, as I could not know which topics from the granted cases were considered detailed enough.

The asylum official's personal interviewing style, their political or ideological biases can greatly affect the level of detail the applicant is able or willing to provide, as well as the decision that is made. As elaborated in the introduction, both laypeople and experts are prone to be biased in their assessments of credibility (Granhag et al., 2005; Maegherman et al., 2018). On the other hand, some interviewers might use more questions and effort to gain the most information possible (Skrifvars et al., 2020), which then causes an unequal situation for the applicants who do not get such interviewers. In my study the interviewer was not one of the coded variables, which was one source of error that could not be controlled in the analysis.

Although the data used for this study was old, it also had ideal timing for researching how a credibility indicator was actually used. The cases in this analysis received their decisions in 2017-2018, and most of the applications were filed in 2015-2016. During 2015 Finland received plenty more asylum applications than the years before, which caused a rush to the asylum system and pressure for politicians. The Finnish Immigration Service had an interest to apply credibility indicators to discredit more people, as Vanto and colleagues (2022) also suggest. Even though the situation with asylum applications might be currently calmer, researching the busy years showcased what can happen when the asylum system gets crowded. The Finnish Immigration Service can choose to be stricter and assess more applicants as not credible when they want to restrict the amount of asylum seekers staying in the country. The refugee situation of today will not be the standard for all the future, and Finland might receive even more asylum applications in the upcoming years due to current wars or climate refugees. Researching a system during a time it was under pressure can give implications of what to expect for the future and how humanely or inhumanely asylum applicants will get treated when global crises get worse.

While the methodology and coding of this study was prone to subjectivity, the study can still explore a new area of how the Finnish Immigration Service works, and how well the guidelines from EU are used and understood in practice. Although credibility indicators, especially level of detail, can be interpreted differently by individual asylum officials, they

should be used according to best practice standards and respectful treatment of asylum seekers. This is why I wanted to look into how the criterion of level of detail was used in the Finnish asylum decisions. To improve the asylum system to be more accurate in its decisions and judgements of the applicants, doing research on its workings is necessary.

5.2 Suggestions for further research

As the statistical analyses in this study were not very robust, replicating the study with a larger amount of data would lend more support to the implications of results here. However, as there seemed to be no statistical connection to actual number of details expressed in the applicant's answers and the assessment of detailedness, it would be interesting to study the same question with a newer dataset. Then it would be possible to compare whether the political climate and the amount of asylum seekers at the time affect the use of credibility indicators, as Vanto and colleagues (2022) suggest. Another area of interest are the interview techniques and their connection to the applicant's level of detail when answering questions. With a new dataset of Finnish asylum decisions, one could study if the methods are in line with best practice and what implications does it have for number of details elicited.

Similar research within the EU, in countries that use the same guidelines and credibility indicators, could bring up possible differences between countries' asylum officials: Is there a country where the criterion of detailedness is used more objectively? This kind of research could have implications for the level of detail -criterion on a larger scale. If it cannot be used objectively, new guidelines could offer concrete methods for measuring it. Although its applicability in practice is debatable, especially in the asylum context, detailedness does have scientific support as a sign of truthfulness (Palena et al., 2021; Verschuere et al., 2021; Vrij et al., 2022). As long as it is being used in credibility assessment, it needs to be applied in an objective and thorough manner. The current way of assessing detailedness does not seem to be functional when applied in practice, as the asylum official cannot be expected to calculate details during the interview or from the transcript either (Vrij et al., 2022). Therefore, the criterion itself could benefit from a new way of applying it, such as estimating the level of detail from the transcript (Vrij et al., 2022). Estimating from the text could be an improved way to assess the detailedness, as it calls for focus and thinking regarding the level of detail. Assessing detailedness by estimating on a Likert scale has some empirical support for connection to actual number of details elicited, but it was mainly not studied on asylum

officials (Vrij et al., 2022). This could be researched more to establish a stronger ground and possibly applied in practice to make level of detail a less subjective credibility indicator.

In order to include all types of asylum applications in the study, a different way to code topics or discarding the use of topics altogether might be needed. In this study I excluded identity-based claims due to the difficulties of coding them similarly to political claims. However, for a full picture of how level of detail is used in credibility assessment, a case-centric approach to counting information units might be more practical than a topic-centric one. On the other hand, counting the details from the entire interview would not reflect the way credibility indicators are intended to be used. The EUAA guidelines (2024) direct the asylum officials to evaluate the credibility of each topic separately, meaning that within an interview some topics can be assessed as sufficiently detailed and some as insufficiently detailed.

The risk of asylum officials focusing on the topics they are interested in could be researched with a statistically stronger dataset. In this study, the number of questions about a topic did have a small effect on the assessment of detailedness, but this finding alone cannot be taken as a fact, considering the statistical weakness and this being only one study. A higher number of questions could be connected to eliciting more details and lead to a topic being assessed as sufficiently detailed. A study looking at these connections could lead to similar findings as mine, if the criterion of detailedness is not used correctly (based on actual details). On the other hand, if the criterion is used correctly, a study with a finding that connects question amount, detail amount and the matching assessment of detail could help establish cut-off points for what counts as enough detail. However, the number of details depends on the extent of the topic itself, in addition to the interviewing technique and individual style of answering. Although cut-off points would be helpful in decision making, they could be extremely difficult to apply in practice due to the variance in number of details.

5.3 Conclusion

The aim of this study was to investigate whether the actual number of details in an asylum seeker's narrative was connected to the Finnish Immigration Service's assessment of detailedness. As no such statistical connection appeared in my analysis, the results imply that the criterion of level of detail was not used objectively and consistently in this sample of cases from Finland. It has been previously suggested that the criterion of detailedness is used without basis (Bodström, 2020). There are also notable challenges in using the criterion in practice, and asylum officials cannot be expected to count details while interviewing someone

(Vrij et al., 2022). While there were limitations that regard caution in my study, it can help establish ground for research regarding the credibility indicators, namely level of detail.

Number of questions and its relationship to number of details as well as assessment of detailedness would be an interesting combination of variables to research with future data. Number of questions and number of details had a correlation in my dataset, which also limits the interpretation of the regression analysis. However, as the number of questions predicted being assessed as sufficiently detailed, it might be that there was an effect hidden behind the question number. Alternatively, if no such effect appears in the future, the asylum officials should be careful when assessing the applicant's detailedness for a topic they chose to ask more questions about. The answers might simply feel like they included more detail, because the matters were discussed a lot in multiple questions. Similarly, topics that were not asked about cannot be assessed as insufficiently detailed, if the applicant was not given a chance to talk about the topic more.

Use of open questions to elicit more information has been supported in previous research, mostly in the criminal interview context and its best practice guidelines, but for asylum interviews too (Brandon et al., 2018; Meissner et al., 2012; Skrifvars, Antfolk et al., 2022; Snook et al., 2012; van Veldhuizen et al., 2018). The results of my analysis also imply that it is a beneficial technique in asylum interviews to gain a more detailed narrative. However, my dataset did not have a great deal of open questions in use nor variation in the proportion of open questions between topics.

In summary, research such as this can, at its best, help improve the asylum system. Although the decisions asylum officials make are not easy, all professionals and institutional practices should receive feedback and keep improving. In my study it appeared that credibility assessments were not conducted according to best practice, which indicates the criterion of level of detail is in need of more applicable and practical customs. Subjectivity and bias in credibility assessments are a risk for unfair decision making, when it is not founded on the applicant's claim and how credibly they can recount it. Asylum interviews and decisions are complex, and both the applicant and the official affect what kind of psychological phenomena is present. However, there are common aspects, such as functioning of memory, communication, biases and expectations that can influence the asylum interview. By investigating what strategies work, what are too subjective, or not based on science of how

people's minds function in situations like these, a more just and humane system could be created.

References

Alien Act, 87 § Asylum. (2024).

https://www.finlex.fi/fi/lainsaadanto/2004/301#chp_6_sec_87v20240666_heading

Allwood, C. M., Ask, K., & Granhag, P. A. (2005). The cognitive interview: Effects on the realism in witnesses' confidence in their free recall. *Psychology, Crime & Law*, *11*(2), 183-198. <https://doi.org/10.1080/10683160512331329943>

Anakwah, N., Horselenberg, R., Hope, L., Amankwah-Poku, M. & van Koppen, P. J. (2020). Cross-cultural differences in eyewitness memory reports. *Applied Cognitive Psychology*, *34*(2), 504-515. <https://doi.org/10.1002/acp.3637>

Bodström, E. (2020). Asylum decisions as performances: Intertextuality in internal credibility assessment. *International Journal of Refugee Law*, *32*(4), 623-644.

<https://doi.org/10.1093/ijrl/eeab001>

Bond C. F. & DePaulo B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*(3), 214-34.

https://doi.org/10.1207/s15327957pspr1003_2

Brandon, S.E., Wells, S., & Seale, C. (2018). Science-based interviewing: Information elicitation. *Journal of Investigative Psychology and Offender Profiling*, *15*(2), 133-148. <https://doi.org/10.1002/jip.1496>

Cameron, H. E. (2010). Refugee status determinations and the limits of memory.

International Journal of Refugee Law, *22*(4), 469-511.

<https://doi.org/10.1093/ijrl/eeq041>

Conway, M. A. (1996). Autobiographical memory. In Bjork, E. L. & Bjork, R.A. (Eds.), *Memory* (pp. 165-194). Academic Press. <https://doi.org/10.1016/B978-012102570-0/50008-2>

Conway, M. A. (2009). Episodic memories. *Neuropsychologia*, *47*(11), 2305-2313.

<https://doi.org/10.1016/j.neuropsychologia.2009.02.003>

Dowd, R., Hunter, J., Liddell, B., McAdam, J., Nickerson, A., & Bryant, R. (2018). Filling gaps and verifying facts: Assumptions and credibility assessment in the Australian Refugee Review Tribunal. *International Journal of Refugee Law*, *30*(1), 71-103.

<https://doi.org/10.1093/ijrl/eeey017>

Dror, I. E., & Charlton, D. (2006). Why experts make errors. *Journal of Forensic Identification*, *56*(4), 600-616.

- European Union Agency for Asylum. (2024). *Practical guide on evidence and risk assessment*. <https://euaa.europa.eu/publications/practical-guide-evidence-and-risk-assessment>
- Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., & Fisher, R. P. (2016). The effect of interpreters on eliciting information, cues to deceit and rapport. *Legal and Criminological Psychology*, 21(2), 286-304. <https://doi.org/10.1111/lcrp.12067>
- Finnish Immigration Service. (2024, November). *Asylum in Finland*. Finnish Immigration Service. <https://migri.fi/en/asylum-in-finland>
- Granhag, P. A., Strömwall, L. A., & Hartwig, M. (2005). Granting asylum or not? Migration board personnel's beliefs about deception. *Journal of Ethnic and Migration Studies*, 31(1), 29-50. <https://doi.org/10.1080/1369183042000305672>
- Gyulai, G. (Ed.). (2013). *Credibility assessment in asylum procedures. A multidisciplinary training manual*. Hungarian Helsinki Committee.
- Herlihy, J., Jobson, L., & Turner, S. (2012). Just tell us what happened to you: Autobiographical memory and seeking asylum. *Applied Cognitive Psychology*, 26(5), 661-676. <https://doi.org/10.1002/acp.2852>
- Herlihy, J. & Turner, S. W. (2009). The psychology of seeking protection. *International Journal of Refugee Law*, 21(2), 171-192. <https://doi.org/10.1093/ijrl/eep004>
- Hope, L., Anakwah, N., Antfolk, J., Brubacher, S. P., Flowe, H., Gabbert, F., Giebels, E., Kanja, W., Korkman, J., Kyo, A., Naka, M., Otgaar, H., Powell, M. B., Selim, H., Skrifvars, J., Sorkpah, I. K., Sowatey, E. A., Steele, L. C., Wells, S. & Anonymous (2022). Urgent issues and prospects at the intersection of culture, memory, and witness interviews: Exploring the challenges for research and practice. *Legal and Criminological Psychology*, 27(1), 1-31. <https://doi.org/10.1111/lcrp.12202>
- Hope, L., Kontogianni, F., Thomas, W. & De La Fuente Vilar, A. (2025). Development and testing of a Time-Critical Questioning protocol for eliciting information in time-sensitive contexts. *Scientific Reports*, 15(1), 14855. <https://doi.org/10.1038/s41598-025-96792-z>
- Kagan, M. (2003). Is truth in the eye of the beholder-objective credibility assessment in refugee status determination. *Georgetown Immigration Law Journal*, 17, 367-415.
- Keselman, O., Cederborg, A.-C., Lamb, M.E. & Dahlström, Ö. (2010). Asylum-seeking minors in interpreter-mediated interviews: what do they say and what happens to their responses? *Child & Family Social Work*, 15(3), 325-334. <https://doi.org/10.1111/j.13652206.2010.00681.x>

- Leal, S., Vrij, A., Deeb, H., Hudson, C., Capuozzo, P. & Fisher, R.P. (2020). Verbal cues to deceit when lying through omitting information. *Legal and Criminological Psychology*, 25(2), 278-294. <https://doi.org/10.1111/lcrp.12180>
- Leal, S., Vrij, A., Vernham, Z., Dalton, G., Jupe, L., Harvey, A., & Nahari, G. (2018). Cross-cultural verbal deception. *Legal and Criminological Psychology*, 23(2), 192-213. <https://doi.org/10.1111/lcrp.12131>
- Maegherman, E., Van Veldhuizen, T., & Horselenberg, R. (2018). Dropping the anchor: The use of plausibility in credibility assessments. *Oxford Monitor of Forced Migration*, 7(2), 37-56.
- Meissner, C.A., Redlich, A.D., Bhatt, S. and Brandon, S. (2012). Interview and interrogation methods and their effects on true and false confessions. *Campbell Systematic Reviews*, 8(1), 1-53. <https://doi.org/10.4073/csr.2012.13>
- Palena, N., Caso, L., Vrij, A., & Nahari, G. (2021). The verifiability approach: A meta-analysis. *Journal of Applied Research in Memory and Cognition*, 10(1), 155-166. <https://doi.org/10.1016/j.jarmac.2020.09.001>
- Selim, H., Helenelund, M., Vanto, J., Skrifvars, J., Alvesalo-Kuusi, A., Pirjatanniemi, E., Korkman, J. & Antfolk, J. (2025). 'How would you define your sexuality?' analyzing the questions asked in official asylum interviews with sexual minorities. *Psychology, Crime & Law*, 1-18. <https://doi.org/10.1080/1068316X.2025.2557965>
- Skrifvars, J., Antfolk, J., van Veldhuizen, T., Sui, V., & Korkman, J. (2022). Eliciting information in official Finnish asylum interviews. *Applied Cognitive Psychology*, 36(3), 508-519. <https://doi.org/10.1002/acp.3936>
- Skrifvars, J., Ilmoni, A., Siegrids, L., Galán, M., Selim, H., Stevens, L., Korkman, J. & Antfolk, J. (2025). Experiences of asylum interviews by asylum officials, interpreters and asylum seekers in Finland. *Journal of Criminal Psychology*, 15(2), 260-280. <https://doi.org/10.1108/JCP-09-2024-0085>
- Skrifvars, J., Korkman, J., Sui, V., van Veldhuizen, T. & Antfolk, J. (2020). An analysis of question style and type in official Finnish asylum interview transcripts. *Journal of Investigative Psychology and Offender Profiling*, 17(3), 333-348. <https://doi.org/10.1002/jip.1557>
- Skrifvars, J., Sui, V., Antfolk, J., van Veldhuizen, T., & Korkman, J. (2022). Psychological assumptions underlying credibility assessments in Finnish asylum determinations. *Nordic Psychology*, 76(1), 55-77. <https://doi.org/10.1080/19012276.2022.2145986>

- Snook, B., Luther, K., Quinlan, H., & Milne, R. (2012). Let 'em talk!: A field study of police questioning practices of suspects and accused persons. *Criminal Justice and Behavior*, 39(10), 1328-1339. <https://doi.org/10.1177/0093854812449216>
- United Nations. (1951). *The 1951 convention relating to the status of refugees*.
- UNHCR. (2025, September 3). *Refugee Data Finder – Key Indicators*. UNHCR, the UN Refugee Agency. <https://www.unhcr.org/refugee-statistics>
- UNHCR. (2026, March 6). *Who is an ‘asylum-seeker’?* UNHCR, the UN Refugee Agency. <https://www.unhcr.org/about-unhcr/who-we-protect/asylum-seekers>
- Vanto, J., Saarikkomäki, E., Alvesalo-Kuusi, A., Lepinkäinen, N., Pirjatanniemi, E., & Lavapuro, J. (2022). Collectivized discretion: Seeking explanations for decreased asylum recognition rates in Finland after Europe’s 2015 “Refugee Crisis”. *International Migration Review*, 56(3), 754-779. <https://doi.org/10.1177/01979183211044096>
- van Veldhuizen, T. S., Maas, R. P., Horselenberg, R., & van Koppen, P. J. (2018). Establishing origin: Analysing the questions asked in asylum interviews. *Psychiatry, Psychology and Law*, 25(2), 283-302. <https://doi.org/10.1080/13218719.2017.1376607>
- Verschuere, B., Bogaard, G., & Meijer, E. H. (2021). Discriminating deceptive from truthful statements using the verifiability approach: A meta-analysis. *Applied Cognitive Psychology*, 35(2), 374-384. <https://doi.org/10.1002/acp.3775>
- Vrij, A., Fisher, R. P., & Leal, S. (2022). How researchers can make verbal lie detection more attractive for practitioners. *Psychiatry, Psychology and Law*, 30(3), 383-396. <https://doi.org/10.1080/13218719.2022.2035842>
- Vrij, A., Hope, L., & Fisher, R. P. (2014). Eliciting reliable information in investigative interviews. *Policy Insights from the Behavioral and Brain Sciences*, 1(1), 129-136. <https://doi.org/10.1177/2372732214548592>
- Wang, Q. (2016). Remembering the self in cultural contexts: A cultural dynamic theory of autobiographical memory. *Memory Studies*, 9(3), 295-304. <https://doi.org/10.1177/1750698016645238>

Appendix

Appendix 1 Coding principles

Variables:

1. Case number

2. Case outcome

Definition: The Immigration Service's decision on the asylum application.

Response-options

1 = asylum granted

2 = other, such as granted secondary protection

3 = asylum rejected

3. Credibility of the topic

Definition: The Immigration Service's opinion whether the topic the applicant has claimed is credible enough to be believed or not.

Response-options

1 = positive credibility finding

2 = negative credibility finding

4. Detail

Definition: Whether the Immigration Service states that the level of detail was insufficient or sufficient, or they do not mention level of detail.

Response-options

1 = level of detail is sufficient

2 = level of detail is insufficient

3 = level of detail is not mentioned

5. Topic

Definition: A topic is an event or fact that is part of the asylum seeker's claim for seeking asylum that the Immigration Service refers to in the decision. In granted decisions, these are listed under the title "Hakemus". Granted cases do not contain any explanations or reasonings for why the application is accepted, otherwise than briefly

listing the topics within the applicant's claim. In rejected cases these are listed under "Tosiseikat". This section contains thorough explanations regarding the topics and their credibility. The coding of topics was done based on the information mentioned in the decisions, not the content of the interviews.

Inclusion criteria

- 1) In rejected cases, any phrase where the Immigration Service describes an event or fact that the applicant has stated in the interview and declares their opinion on, is used as a topic. These phrases are such as: "Migri hyväksyy/ei hyväksy tosiseikkana..." and "Migri katsoo...", where the name of the organization is present.
- 2) In granted cases, use full sentences describing the asylum claims as topics, because the text is short and does not contain information of the Immigration Service's reasoning. They must also be related to the reason for granting asylum (mentioned under "Perustelut", e.g., political opinion), The full sentence can be divided to different topics only when there are clearly separate topics listed in one sentence.

Exclusion criteria

- 1) Exclude if there are interviews missing regarding the topic, meaning when there are multiple interviews where the topic is addressed, but not all are in our materials.
- 2) Exclude if the topic is a shared experience between family members and not all interviews are in our materials. Other family members should be referred to in the decision text, for example by using a plural form or assessing credibility based on similarity of their answers.
- 3) Common conclusions (of topics already mentioned in the decision text before) are not a separate topic. Example of a common conclusion: "Therefore Migri does not accept that the militias or the government cause you any serious threat upon return.", if concern regarding these two was already elaborated previously. Example of a conclusion that brings new information: "Migri considers your fear is based on your own assumptions, not on your religion and father's background", because father's background is only mentioned here.

- 4) Exclude if the topic is disregarded based on previous topics not being accepted.
Example of this: “Migri does not accept that you’re being stalked, because Migri did not accept the earlier events or the house search.”
- 5) If a topic is not investigated in the asylum interview despite being mentioned in the decision text, exclude it. These can be topics that are based on physical evidence or disclosed in a previous interview at the border.
- 6) If a topic is of similar theme to another topic and the questions are heavily overlapped, combine these two to a single topic. Heavily overlapped = when all or almost all question numbers are used in both.

6. Question-answer pair number

Definition: Number of the question in the asylum interview.

If the topic is mentioned either in the question or answer, write it down. This applies also when the topic is simply mentioned: even if the topic is not elaborated or only previous information is repeated, include the question-answer pair. Same question-answer pairs can be included in multiple topics.

Consider before including:

- 1) If the relation to the topic is unclear, see if what is talked about is regarded as a separate topic evaluated by the Immigration Service (in the decision). If it is, do not include it for this topic. Example: If there are separate topics, such as “you were imprisoned by group x” and “you were tortured in prison by group x”, questions about torture will only be under the second topic, but being taken by the group and descriptions of prison will be under the first topic.
- 2) If a question is open or does not refer to a specific topic, if the context (such as previous or following questions clarifying the topic) implies it is related to the event in question, include it. Example: If a topic is “you are afraid of group x”, question-answer-pairs “What are you afraid will happen to you, if you return? – I will die.” and “Who would kill you? – Group x.” will both go under the topic. This is because the second one explains how the topic (group x) is related to the first one (fear of death).

Exclusion criteria:

- 1) The topic is mentioned only as a time indicator for some other topic. Example: “I was shot at two days after my sister died.” would be included in the topic of being shot at, but not the topic of the sister’s death.
- 2) It is unclear which topic is discussed and there is no new information presented. Example: “What are you afraid will happen to you, if you return? – I will die.” without any clarification why, by whom or anything to relate it to a topic, this should be excluded.

7. Notes