Executive functioning and its role in high impact chronic pain. Building a causal model using Directed Acyclic Graphs.

Annick L. De Paepe[1], Anna Gibby[2], Laura Oporto Lisboa[3], Beate Ehrhardt[3], Matthew Nunes[7], Emma Fisher[2], Edmund Keogh[2], Christopher Eccleston[1,2,4], Charlotte Woolley[5], John McBeth[5,6] & Geert Crombez[1] on behalf of the CHIPP group

[1] Faculty of Psychology and Educational Sciences, Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium.

[2] Centre for Pain Research, University of Bath, UK

[3] Institute for Mathematical Innovation, University of Bath, Bath, UK

[4] Department of Psychology, The University of Helsinki, Finland

[5] Faculty of Biology, Medicine and Health, School of Biological Sciences, University of Manchester, UK

[6] Faculty of Medicine, School of Primary Care, Population Sciences and Medical Education, University of Southampton, UK

[7]Department of Mathematical Sciences, University of Bath, Bath, UK

1 Introduction

Chronic pain, as well as its impact, can differ between individuals and over time, and is influenced by a myriad of biological, psychological and social factors [13,32]. Eccleston and colleagues [11] provided an overview of different pain states and their transitions. An important scientific and clinical question is which factors influence an individual's transition from one pain state to another, or indeed result in the lack of transition. This is a causal question and challenging to answer, because of the inability to experimentally manipulate many relevant factors. Instead, we rely on natural (uncontrolled) variations that occur within individuals and study how these variations differentially impact transitions in chronic pain. The role of confounding factors is a well-known problem in observational and longitudinal studies [17]. Researchers argue that it is impossible to infer causality from non-experimental designs and have, for that reason, avoided causal thinking. Several papers highlighted the dangers inherent in this avoidance, including the risk of bias and misinterpretation of effects [5,7,18,49,50].

Causality cannot be inferred from data alone (i.e. bottom-up approach). As Judea Pearl aptly states, "Data are profoundly dumb" [37]. Data might show that people with poorer executive functioning are more likely to develop chronic pain, but we don't know its actual drivers. For example, it could be that those with poorer executive functioning are older and would have developed chronic pain regardless of their cognitive abilities. To help reinstate causal thinking, Directed Acyclic Graphs (DAGs) have been promoted as an innovative tool to represent the assumptions (based on domain knowledge) about the causal relationships between variables in a graphical model [10,36,41,45]. A DAG thereby empowers researchers to engage in a constructive discussion about underlying relationships. They help identify the smallest set of variables needed to be adjusted for to remove confounding bias in estimating the causal effect of an exposure on an outcome ('minimal sufficient adjustment set') (Figure 1). The DAG in Figure 1 illustrates that 'age' is a confounder, indicating that we need to control for it when examining the causal effect of executive functioning on chronic pain.

[Insert Figure 1 here]

Despite their adoption in many scientific areas, there are currently few examples of DAGs being used in pain research [28,42,46,51]. The main barrier to adopting DAGs is lack of knowledge about their use and construction [4]. Although there is guidance on the mathematical underpinnings of causal diagrams, few articles give practical guidance to researchers on how to extract domain knowledge and subsequently construct a DAG [39]. Further, when DAGs have been used it is common for researchers to not report their method, and critically whether or how the DAG guided the analysis [45].

In this paper, we provide a worked example with explanations on how to build a DAG based on domain knowledge obtained from the literature, researchers, and individuals with lived experience. The use case is the putative effect of executive function on the transitions between chronic pain states.

2 Methods

DAGs are constructed based on domain-relevant knowledge. Domain knowledge in this example was obtained from three different sources: the literature, a group of researchers and a group of individuals with lived experience (ILE) of chronic pain. To retrieve and reconcile information from the different sources of domain knowledge we followed four phases based upon Rodrigues and colleagues [40] and complemented by the steps of DAG development outlined by Poppe et al. [39]: (1) brainstorming phase: gather information from all individuals in the workshop separately; (2) refinement phase: reach consensus amongst the group members; (3) exposition phase: obtain feedback from the researcher group on the DAG developed by the ILE group and the DAG developed based on the literature; and (4) reconciliation phase: combine features of all DAGs into a final DAG.

The brainstorming and refinement phase were conducted by the researcher group and the ILE group in separate workshops. Therefore, in the description below, these two phases are taken together. Two separate workshops were organised for researchers and the ILE group to ensure that everyone felt free to communicate their ideas without boundaries. A third follow-up workshop was organised for the researcher group in which the exposition and reconciliation phase was done. An overview of the different phases followed during the study are given in Figure 2.

[Insert Figure 2 here]

It is important to note that we prospectively planned to use data from the UK Biobank (https://www.ukbiobank.ac.uk/) to answer our research question, resulting in an iterative process. The availability of factors measured by the UK Biobank has influenced the specification of our research question and the definition of our exposure and outcome (see step 1 in section 2.1.1.2). However, we were not guided by the UK Biobank when thinking about potential confounders or the relationships between variables. This was done for two reasons. First, it is important to consider all potential confounders in a DAG, even if these were not measured in the available dataset. Unmeasured variables can also be represented in a DAG and several methods have been suggested to deal with unmeasured confounders [2]. Second, our aim is to provide a DAG with a more general interpretability aiding those investigating related research questions.

2.1 Workshop 1 with researchers

2.1.1 Participants

Eight researchers took part in a DAG development workshop. The group of researchers consisted of a mix of people with domain knowledge (ADP, GC, CE, EF, EK, AG) and researchers with a strong background in statistics and methodology (ADP, MN, LOL). Details with regards to the affiliation and background of the researchers can be found in Supplementary Information (SI-Table 1).

2.1.2 Procedure

An in-person half-day introduction to causal DAGs was given by ADP to the researchers in January 2023. This was followed by a one hour online ‘refresher’ course in August 2023. Moreover, several introductory articles on DAGs were shared with the researcher group. We considered it unnecessary for domain experts to fully understand the mathematical graph theory underpinning DAGs, but the rationale of their use in research was deemed necessary to be able to contribute in a meaningful way to the development of the DAG. Next, an in-person DAG development workshop of two half-days was held in October 2023. The purpose of this workshop was to create a DAG for one specific research question. Three steps were followed to construct the DAGs as outlined by a recent scoping review on guidelines for DAG development [39]: (1) specify the

research question and define the exposure and outcome; (2) add common causes and arrange temporally; (3) draw forward/directed arrows indicating the assumed causal relationship between two variables. On the first day we focused on step 1 and 2. The second day was devoted to finishing steps 2 and 3. During the workshop a shared MIRO board (https://miro.com/nl/, [33]) was used to facilitate the discussion.

2.2 Workshop 2 with individuals with lived experience

2.2.1 Participants

Seven people with lived experience of chronic pain (6 women, 1 man), took part in a workshop that was facilitated by two researchers (AG and ADP). Six of the participants had a formal diagnosis of one or more pain conditions. The conditions reported were fibromyalgia (N=4), osteoarthritis (N=2), neuropathy (N=1) or another unspecified condition (N=2). Pain was reported in back, muscles/tendons/joints, head, or all over. Participants ranged from 25-70 years old, five individuals reported being from a lower income and two reported belonging to an ethnic minority group. Employment status included working (N=3), primary carer (N=2), student (N=1), and retired (N=1). The Guidance for Reporting Involvement of Patients and the Public Short Form (GRIPP2-SF) reporting checklist [44] was completed and can be found in SI (SI-Table 2).

2.2.2 Procedure

A 2-hour online meeting was held in November 2023. During this meeting a short introduction was given about the aims of the research project in non-technical terms by one of the facilitators. The independent and dependent variables were introduced as ‘thinking patterns/cognitive factors’ and ‘chronic pain experience’ as a lay substitute for cognitive factors and impact of chronic pain respectively and they were briefly explained. Also, ‘transitions of impact in pain’ were explained briefly and we explained that within this study we focused specifically upon the maintenance of high impact chronic pain. Finally, we introduced the main question that we would focus upon during the workshop as follows: “We predict that thinking ability can cause chronic pain status over time. We need help thinking about what other factors might impact this relationship”. An example was given to clarify (depression as a common cause of thinking ability and chronic pain status).

During the workshop with the ILE group we planned to focus upon ‘Step 2’ of the researcher workshop, namely ‘add common causes and arrange temporally’. However, because of the limited time available we were not able to have patients’ input on the temporal ordering of the factors. This step was done by the facilitators of the workshop (AG and ADP) and was subsequently discussed with the co-authors. During the workshop a shared MIRO board (https://miro.com/nl/, [33]) was used to facilitate the discussion.

2.3 DAG based on literature

Next, we identified factors in the literature. Papers were considered for inclusion based on the following criteria:

1. The study was a prospective, longitudinal design
2. Pain status was measured at both baseline and follow up
3. Executive function was explored as a predictor of pain

Papers that met these inclusion criteria were identified in several ways. First, based on the 734 included papers for another review (search conducted in March 2022) with similar but broader inclusion criteria, from colleagues at the Advanced Pain Discovery Platform (ADPD) Consortium to Research Individual, Interpersonal and Social Influences in Pain (CRIISP, https://criisp.uk/). These papers were scanned to identify papers that matched our more focused inclusion criteria. Second, a brief search of the literature was done in PubMed and Google Scholar using key terms relevant to the inclusion criteria. The terms used were ‘executive function & pain’, ‘executive function & chronic pain’, ‘cognitive factors & chronic pain’ and ‘Trail Making Task and Pain’, as these were most relevant to the main research question. This search was limited in terms of timeframe (from 2021 – present, since this period was not included in the review conducted by the CRIISP consortium) and inscape (papers had to be longitudinal and include the impact of executive function on pain outcomes to identify confounders considered for both). The researcher group involved in workshop 1 was then asked to identify any papers that may have been missed.

For each relevant study, the variables that were controlled for in these studies were extracted by AG and ADP and a separate DAG was created for each study. Subsequently, all variables identified were temporally ordered and integrated in one overarching DAG.

2.4 Workshop 3 with researchers (follow-up)

2.4.1 Participants

Seven of the eight researchers who participated in the DAG development workshop took part in an follow-up online workshop for the exposition and reconciliation phase in March 2024.

2.4.2 Procedure

The workshop focused on the exposition and reconciliation phase. The background and aims of the study and a summary of the different steps that had already been taken in the DAG development process were presented. Before the workshop, AG and ADP had compared the three DAGs and extracted those factors that were identified by the ILE group and/or the literature but were not present in the researcher DAG. During the workshop, the research team was presented with these new factors and the corresponding reasoning of the ILE group or the literature (exposition). The team discussed whether to include these factors either as common causes or as mediators (reconciliation).

3 Results

3.1 Workshop 1 with researchers

3.1.1 Step 1: Specify research questions and define the exposure and the outcome

In the ***brainstorm phase***, all researchers were asked to think independently about two questions: (1) “If we had all the data that we wanted, what is/are the most interesting research question(s) that you would like to answer?” and (2) “Given the data that we have, which research questions are feasible to answer?”. Researchers were asked to type their thoughts on virtual post-it-notes that they could add to the shared MIRO board. Results of this phase are presented in Supplementary Information (SI-Figure 1).

In the ***refinement phase*** a consensus was reached among the researchers to focus upon ‘Executive functioning’ as exposure and on the ‘Maintenance of high impact chronic pain’ as outcome.

Exposure. Executive functioning involves higher-order cognitive strategies that control stimulus information and subsequent behaviour identified by tests of inhibition, shifting, and working memory updating [34]. Specifically,

inhibition involves suppression of prepotent responses (i.e. response inhibition) and selectively attending to motivationally relevant stimuli, and not attending to irrelevant stimuli. Shifting involves the ability to switch between mental sets and rules, and updating involves monitoring information in working memory with addition, suppression, and deletion. Within the UK Biobank executive functioning was measured using two different tasks: 'Trail making test A and B' and 'Tower of London'. Different components of executive functioning are measured by the two tasks (see Table 1). We decided to focus upon the trail making test part A and B as independent variable because this test measures components of executive functioning that we considered most relevant in the context of the maintenance of chronic high impact pain. Moreover, we decided to work with the difference score between part B and A, as this has been shown to provide a relatively pure indicator of executive control abilities, predominantly measuring switch-cost [43].

Executive functioning was not measured at the first measurement period of UK Biobank (2006-2010). Additionally, there is evidence of inconsistent validity of outcomes for the cognitive battery in the 2013 follow up [30], as well as a limited sample size of approximately 20,000 participants (which represents only 4% of the total sample). Executive function was also measured in 2014 and 2019. Although these measurement periods started in 2014 and 2019 respectively, not every participant completed the assessments in that year. In the cohort that will be used for the analyses, we will ensure that there is at least a 6 month difference between the two measurement points. It is parsimonious to consider these simply as two temporally separated time points. Hereafter we refer to these as Time 0 (2006-2010), Time 1 (starting in 2014) and Time 2 (starting in 2019).

	Components of executive function measured[1]	**Outcome**
Trail making test part A and B	Cognitive control Serial processing Visual search	Time (in deciseconds) to complete part A and B
Tower of London (imagine)	Strategy Speed of planning	Number of items answered correctly in three minutes

Table 1. Details of the two tasks measuring executive functioning included in the UK Biobank.
[1]Components of executive functioning are based on the cognitive atlas (https://www.cognitiveatlas.org/)

Outcome. We were interested in the maintenance of high impact chronic pain. To investigate this, participants who had high impact chronic pain at baseline and follow-up were compared to participants who had high impact chronic pain at the baseline, but whose functioning improved at follow-up to chronic pain with a low impact. Definitions of chronic pain, low and high impact chronic pain and the transitions in pain states are presented in Table 2.

	Definition
Chronic pain	Pain that has lasted for three months or more
High impact chronic pain	Chronic pain with a high impact on major self-care, occupational or social activity restrictions.
Low impact chronic pain	Chronic pain with a low impact on major self-care, occupational or social functioning.
Onset of chronic pain	Change from no pain to chronic pain with a low or high impact
Maintenance of chronic pain	No change in chronic pain state.
Worsening of chronic pain	Change from chronic pain with a low impact to a high impact
Improvement of chronic pain	Change from chronic pain with a high impact to a low impact
Recovery of chronic pain	Change from chronic pain with a low impact to no pain

Table 2. Definitions of chronic pain, high and low impact pain and transitions in pain states taken from [10]

Research question. To correctly specify a research question, it is important that it consists of the following parts: (1) estimand: the effect the researcher wishes to estimate from the data; (2) exposure of interest: the values or levels of the exposure at which one expects a change in the outcome (e.g. clearly defining ‘good’ versus ‘poor’ executive control); (3) outcome: it is important to choose an appropriate time window to study the effect of the exposure on the outcome. One has to ensure that the outcome is separated in time from the exposure, because a cause should precede the effect. If the exposure is highly variable, it may be necessary to choose a smaller time window. Alternately, if it takes time before the exposure can have an effect on the outcome, then it is necessary to choose a larger time window; and (4) population: a DAG is only valid within a specific population or context. If one wants to investigate the same research question in a different population or context (e.g. different country), one has to reassess whether all relationships specified within the DAG are still valid.

Our research question is: “What is the average causal effect of low executive functioning (a score below the population norm on the trail making

test) compared to normal or higher executive functioning (a score equal to or higher than the population norm) at Time 1 on the maintenance of high impact chronic pain at Time 2 in those with High Impact Chronic Pain at Time 1 (Table 3).

Estimand	Average causal effect: the average change in the outcome that can be attributed to a difference in the exposure.
Exposure	A score below the population norm on the trail making test (i.e. poor executive functioning) as opposed to a score equal to or higher than the population norm (i.e. normal executive functioning) at baseline.
Outcome	The maintenance of high impact chronic pain at Time 2
Time difference between baseline and follow-up	At least 6 months (T1-T2)
Population	Respondents of the UK Biobank who have chronic pain of high impact at Time 1

Table 3. Different components of the research question.

3.1.2 Step 2: Add common causes (or confounders) and arrange temporally

In the ***brainstorm phase*** researchers thought individually about possible common causes of the exposure and the outcome (also called 'confounders'). They were asked to order these common causes temporally, from left (variables that were fixed early in the life of the participants) to right (variables that are time-variant or were only fixed more recently). Researchers were asked to type their thoughts on virtual post-its and add these to the shared MIRO board. Results of this brainstorming phase are presented in SI-Figure 2. Researchers spontaneously brought up some of the mediating pathways between the exposure and the outcome. Because we are interested in the average causal effect, controlling for mediators could lead to an underestimation of the influence of executive functioning on the maintenance of high impact chronic pain. Therefore, we decided to also think about the most important mechanisms by which executive functioning could influence the maintenance of high impact chronic pain, to make sure that we do not mistakenly take these variables into account as possible confounders.

In the ***refinement phase*** a group discussion evolved around the common causes and the mediators of our research question of interest. First, variables that were essentially the same but were formulated differently were merged together. For example, ‘early pain experiences’ and ‘pain exposure during early years’ were combined to ‘early pain experiences’. Second, we thought about which groups of variables that occurred at similar time points and had similar relationships to the other variables that were specified in the DAG. These variables were grouped together in what is sometimes called ‘super-nodes’ [45]. For example, age, socio-economic status, education, profession, country of origin and first language were taken together in the super-node ‘Demographic factors’. Third, we discussed the most important mediators between executive functioning and the maintenance of chronic high impact pain. The result of this phase can be found in SI-Figure 3.

Consensus was reached about most of the common causes and mediators. There was uncertainty about whether to include ‘Lifestyle factors’ as a mediator or a confounder. We decided to define it more precisely as health-related behaviours (diet, physical activity, treatment adherence, smoking) and to add it as a mediator. Nevertheless, we decided to conduct a sensitivity analysis, looking at the difference in the average causal effect when considering it as a confounder as opposed to a mediator.

3.1.3 Step 3: Draw forward arrows

All the forward arrows (i.e. arrows going towards variables to the right in the DAG) were added to the DAG and a group discussion evolved around which arrows to leave out of the DAG. If there was uncertainty about whether an arrow should be removed, we decided to keep the arrow in the DAG, as removing an arrow is a stronger assumption of no association than keeping the arrow in [45]. Results of this phase are shown in SI-Table 3.

We decided to leave the genetic factors and environmental factors out of the DAG, because we considered the direct impact of these variables on executive functioning and impact status of chronic pain at follow up to be negligible.

3.1.4 Elaborating the definitions of the constructs included in the DAG

To be sure that all constructs entered in the DAG were in the appropriate place, it is important to be explicit about definitions of the constructs we adhere to. Therefore, after the two-day DAG development workshop, clear definitions were added for the constructs present in SI-Figure 3. Definitions were as far as possible based on existing definitions and were discussed within the group of researchers until a consensus was reached. By clearly defining all the concepts within the DAG, we became aware of some overlap between the constructs that were present within the DAG (e.g. pain anxiety and pain-related fear). Moreover, we decided to omit some factors at this point, because when reconsidering them we judged that a direct relationship between the construct and either the independent or dependent variable was unlikely (e.g. after consulting with experts on pain-related stigma, stigma was judged unlikely to have a strong association with executive functioning). An overview of the changes made at this point are provided in Supplementary Information (SI-Table 4). The final definitions of the included constructs can also be found in Supplementary Information (SI-Table 5). The resulting DAG for the researcher group is shown in Figure 3.

[Insert Figure 3 here]

3.2 Workshop with people with lived experiences

3.2.1 Step 1: Generation of common causes

In the ***brainstorm*** phase, the ILE group was asked to generate any ‘thinking factors’ they believed influenced their ‘chronic pain experience’. They were asked to type down their thoughts on virtual post-its and add it to the shared MIRO board. The 60 factors that were generated can be found in Supplementary Information (SI-Figure 4).

During the ***refinement*** *phase* we asked the ILE group to expand on factors less clear to the facilitators. Next, the facilitators grouped the factors into ‘super-nodes’, similar to the researcher group (see SI-Table 6). Finally, each one on the ILE group was asked to outline the three factors that they felt were most important in affecting their pain experience. The factors mentioned included lack of care consistency (N=1), trauma (N=2), invisible illness (N=1), weather (N=2),

stress, pacing (N=2), mental health (N=1), sleep (N=3), being validated (N=1), brain fog (N=1), clarity (N=1), expectation (N=1), dissociation (N=1) and energy (N=1).

3.2.2 Step 2: Arrange the factors temporally

After the meeting, the facilitators put the common causes identified by the ILE group in temporal order. The results are shown in Figure 4.

[Insert Figure 4 here]

3.3 DAG based on literature

Five papers were identified as matching the inclusion criteria [3,14,15,34,48]. One to eight confounders were considered in the analyses of the studies. None of the studies reported to have constructed a DAG to answer their research question. An overview of all the common causes and potential mediators identified for the research question based on the literature are given in Figure 5. It is important to note that all included studies focused on the development of chronic pain, whereas our research question focuses on the maintenance of chronic pain. Moreover, four out of the five studies had only pain severity and the presence of chronic pain as outcome, and one study had pain interference as secondary outcome. In contrast our focus is on the impact of chronic pain. Details of the studies and DAGs constructed based on the information given in the papers are outlined in Supplementary Information (SI-Table 7).

[Insert Figure 5 here]

3.4 Workshop 3 with researchers (follow-up)

There was a lot of overlap in the common causes suggested by the researcher group and ILE group. Based on the suggestions of the ILE group, we decided to add ‘lack of care consistency’ and ‘pacing’ as mediators, and ‘medication side-effects’ and ‘trauma later in life’ as common causes. Factors that were suggested by the ILE panel that were not included in the final DAG were contextual factors (e.g. treatment availability, the weather), social factors (e.g. the role of pets, stigma, social support), cognitive factors (e.g. memory, brain fog) and genetic factors.

Based on the DAG constructed based on the literature it was decided to add ‘income quartile’, ‘employment’ and ‘profession’ as more precise indicators of

SES and to include ‘sex’ under demographic factors. Other variables such as certain coping strategies (e.g. religious coping, passive coping) or personality traits were not included.

Extensive reasoning for (not) including the variables suggested by the ILE panel and in the literature are reported in Supplementary Information (SI-Table 8). The final DAG is depicted in Figure 6.

[Insert Figure 6 here]

4 Discussion

In pain research, causal questions often arise, even when working with observational data. A key distinction between a predictive approach and an approach aimed at making causal inferences is the necessity of incorporating domain knowledge in the latter. To make valid causal inferences, it is essential to carefully consider variables that influence both the exposure and the outcome (i.e. confounders) so that the exposure’s parameter estimates can be interpreted causally [43,48]. The domain knowledge needed to draw causal inferences can be summarised and visualised in Directed Acyclic Graphs (DAGs). In this paper, we outlined the steps involved in constructing a DAG based on domain knowledge for the putative effect of executive function on the transitions between chronic pain states. The end result is a DAG that can provide a framework for guiding future research on the role of executive functioning on pain.

In this paper, we demonstrated a procedure for DAG development, based on Rodrigues et al. [40] and Poppe et al. [39], for a relevant and timely research question in pain research. In our DAG development process, considerable time was spent on formulating the research question and on clearly defining the exposure and the outcome. Causal research questions are not easy to formulate. A helpful approach is to define the causal effect as the effect that would have been observed in a hypothetical trial. In other words, you can look at causal inference from observational data as an attempt to emulate a target trial [9,23,24]. This includes clearly specifying the population under study (or the eligibility criteria), the exposure (or treatment strategy), the outcome and the causal contrast of interest (or estimand). Thinking about a research question in this way ensures clear specificity of the causal contrast of interest (e.g. the average treatment effect), both the ‘active treatment’ condition (e.g. poor

executive functioning) and the comparator (e.g. good executive functioning), its operationalisation (e.g. a score above the population mean versus below the population mean on the trail making test) and the time points at which the outcome will be evaluated (e.g. 6 months). This ensures an appropriate time window between the exposure and the outcome, where they are clearly temporally separated, and it is likely for the exposure to have an effect on the outcome.

In observational studies, causal questions will require adjustment for confounding variables to achieve comparability (or exchangeability) between groups with different levels of exposure. Several guidelines help determine whether a variable qualifies as a confounder [47]. Here, we followed the ‘common cause’ approach, including all pre-exposure variables that influence both exposure and outcome [16]. Researchers, individuals with lived experience (ILE) and the literature were consulted to identify the most important confounders and put them in temporal order. While there was overlap among sources, each provided unique variables, highlighting the value of diverse perspectives. Ultimately, 18 confounders were included in the final DAG—more than the minimal variables used in similar studies [3,14,15,34,48]. Despite thorough efforts, it is possible that some confounders were missed, as no DAG perfectly represents reality. The major benefit of including DAGs in research is their transparency, explicitly outlining the assumptions underpinning the estimated causal effect. This aligns with the principles of open science [8] and establishes a firm groundwork for scientific discussion. In many statistics courses the mantra "Correlation does not imply causation" has led to avoiding causal language and thinking. Traditionally, causal inferences are only made from randomized experiments, not from observational data. However, many situations don’t allow for experimental manipulation, leading researchers to avoid causal language in observational studies but make implicit causal inferences in their discussions. For example, Giusti et al. [15] discussed executive functions ‘shaping’ pain outcomes, implying causality, even in a predictive context. Using a predictive approach for causal claims is a common mistake in many studies, leading to misinterpretation of results and even erroneous conclusions [18,22]. Instead, we should be transparent about the causal goal of our analysis [19].

The ‘common cause’ principle identifies confounders as variables that *happen prior to the exposure* and that also have an influence on the outcome.

This timing is crucial to avoid mistaking mediators for confounders, as adjusting for mediators can induce bias. A mediator explains part of the process through which an exposure affects the outcome, so adjusting for it would remove some of the effect of interest. However, in practice, we are often confronted with datasets in which the exposure and potential confounders were all measured at the same timepoint. While it is obvious for some variables that they occurred or 'crystallised' before the exposure (e.g. early pain experiences), it is less clear for others (e.g. physical activity). This complicates decisions on whether to classify a variable as a confounder or a mediator. During the workshops we encouraged participants to also think about potential mediators, to ensure that we would not adjust for these variables. Sometimes there was uncertainty about whether a variable was more important as a confounder or as a mediator. We kept track of this information and will conduct sensitivity analyses to determine the influence of adjusting or not adjusting for these variables on the effect estimate [45].

Apart from clearly defining the exposure and the outcome, it is also important to clearly define confounders and mediators. This was a useful exercise in our case. We were able to merge or exclude some of the variables due to content overlap or to changes in their temporal order. Being explicit and transparent about causal assumptions also requires to be explicit and transparent about their definitions. There is increasing awareness that some often used concepts are poorly defined [7,12,38] and efforts have been made to provide an overview of definitions of commonly used concepts within pain research and to offer guidance to do better [6,25,35].

Of further note, DAGs can also be used to include selection [21,29] and measurement bias [20]. Selection and measurement bias are subject to the specific study design and measurement instruments used. The focus of this paper was to develop a conceptual DAG, which may serve as a starting point for other researchers to design studies or inform data-analysis. In a next step, this conceptual DAG will be taken forward and used to inform an empirical analysis. At that stage the variables in the conceptual DAG must be matched on the actual variables at hand in the dataset. It is important to ensure that the variables are operationalized in a way that matches the definition of the constructs and that minimises measurement bias. If some of the variables are not available within the dataset, they can be included in the DAG as 'unmeasured' and one could for example conduct bias analysis for uncontrolled confounding [2,31]. Also, the way

in which participants were included in the sample should be represented in the DAG. For example, in the UK Biobank there is evidence that the individuals included are more likely to be older, female, and of higher socio-economic status than the general UK population. This could lead to spurious associations (e.g. older individuals in the UK Biobank tend to be in better health). If researchers are aware of this there are ways to, at least partially, remediate [1].

Finally, we presented one possible approach to include causal thinking in pain science. Other approaches exist that may complement this approach. For example, machine-learning tools might help researchers to improve confounding adjustment compared to the traditional approach [26]. This may be especially relevant in the context of large clinical databases that are not specifically collected for the specific research questions under investigation, because many important confounding factors will not be measured in these data sources [52]. It should be noted that domain knowledge remains crucial to causal inference. As it has been pointed out, including numerous possible confounders without theoretical background knowledge to guide the choice increases the likelihood of inducing bias [27].

This study has some limitations. First, although we incorporated the perspectives of researchers and ILE in our DAG, we did not involve other perspectives, such as clinicians. Second, due to lack of time and resources not all three steps of the DAG development process could be completed with the ILE group. Third, for the DAG based on the literature, we had to rely on literature for related research questions that were similar, but not identical to our own research question. However, we discussed each of the variables in the light of our own research question and included them only if we deemed it likely that they would confound.

In this paper, we aimed to highlight the importance of explicitly defining the causal goal of an analysis, even when using observational data. We demonstrated a structured procedure for developing a DAG, by integrating domain knowledge from researchers, individuals with lived experience and existing literature. The resulting DAG provides a valuable framework for guiding future research on the role of executive functioning in pain and it underscores the broader potential of using DAGs to improve causal inference in pain research.

5 Acknowledgements

We would like to extend our deepest gratitude to the individuals with lived experience who generously contributed their time, insights and personal narrative to this work. This work was supported by a joint and equal investment from UKRI [grant numbers MR/W004151/1, MR/W026872/1 and MR/W002388/1] and the charity Versus Arthritis [grant number 22891] through the Advanced Pain Discovery Platform (APDP) initiative. For UKRI, the initiative is led by the Medical Research Council (MRC), with support from the Biotechnology and Biological Sciences Research Council (BBSRC) and the Economic and Social Research Council (ESRC). GC, ADP, CE, EF & EK are part of the research scientific network grant PAIN, funded by the FWO (W001721N). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Edmund Keogh reports unrelated consultancy services via the University of Bath to Reckitt Benckiser Health Limited.

6 Authorship statement

The following are members of CHIPP (Chronic High Impact Pain Project): Payam Amani, Marcus Beasley, Milica Bucknall, Claire Burton, Geert Crombez, Annick De Paepe, Christopher Eccleston, Beate Ehrhardt, Emma Fisher, Anna Gibby, Sarah Harrisson, Edmund Keogh, Laura Oporto Lisboa, John MacBeth, Gary MacFarlane, Barbara Nicholl, Emma Parry, Kerry Summerfield, Danielle Van der Windt, Charlotte Woolley.

7 References

[1] van Alten S, Domingue BW, Faul J, Galama T, Marees AT. Reweighting UK Biobank corrects for pervasive selection bias due to volunteering. Int J Epidemiol 2024;53.

[2] Arah OA. Bias Analysis for Uncontrolled Confounding in the Health Sciences. Annu Rev Public Health 2017;38:23–38. doi:10.1146/annurev-publhealth.

[3] Attal N, Masselin-Dubois A, Martinez V, Jayr C, Albi A, Fermanian J, Bouhassira D, Baudic S. Does cognitive functioning predict chronic pain? Results from a prospective surgical cohort. Brain 2014;137:904–917.

[4] Barnard-Mayers R, Childs E, Corlin L, Caniglia EC, Fox MP, Donnelly JP, Murray EJ. Assessing knowledge, attitudes, and practices towards causal directed acyclic graphs: a qualitative research project. Eur J Epidemiol 2021;36:659–667.

[5] Van Cauwenberg J, De Paepe A, Poppe L. Lost without a cause: time to embrace causal thinking using Directed Acyclic Graphs (DAGs). International Journal of Behavioral Nutrition and Physical Activity 2023;20:145. doi:10.1186/s12966-023-01545-8.

[6] Crombez G, De Paepe AL, Veirman E, Eccleston C, Verleysen G, Van Ryckeghem DML. Let’s talk about pain catastrophizing measures: An item content analysis. PeerJ 2020;2020.

[7] Crombez G, Veirman E, Van Ryckeghem D, Scott W, De Paepe A. The effect of psychological factors on pain outcomes: Lessons learned for the next generation of research. Pain Rep 2023;8:E1112.

[8] Crüwell S, Van Doorn J, Etz A, Makel MC, Moshontz H, Niebaum JC, Orben A, Parsons S, Schulte-Mecklenbeck M. Seven Easy Steps to Open Science: An Annotated Reading List. Zeitschrift fur Psychologie / Journal of Psychology 2019;227:237–248.

[9] Dang LE, Dang LE, Gruber S, Lee H, Dahabreh IJ, Stuart EA, Williamson BD, Wyss R, Díaz I, Ghosh D, Kıcıman E, Alemayehu D, Hoffman KL, Vossen CY, Huml RA, Ravn H, Kvist K, Pratley R, Shih M-C, Pennello G, Martin D, Waddy SP, Barr CE, Akacha M, Buse JB, Van Der Laan M, Petersen M. A causal roadmap for generating high-quality real-world evidence. J Clin Transl Sci 2023;7:1–12. doi:10.1017/cts.2023.635.

[10] Digitale JC, Martin JN, Glymour MM. Tutorial on directed acyclic graphs. J Clin Epidemiol 2022;142:264–267. doi:10.1016/j.jclinepi.2021.08.001.

[11] Eccleston C, Begley E, Birkinshaw H, Choy E, Crombez G, Fisher E, Gibby A, Gooberman-Hill R, Grieve S, Guest A, Jordan A, Lilywhite A, Macfarlane GJ, Mccabe C, Mcbeth J, Pickering AE, Pincus T, Sallis HM, Stone S, Van Der Windt D, Vitali D, Wainwright E, Wilkinson C, De AC, Williams C, Zeyen A, Keogh E. The establishment, maintenance, and adaptation of high-and low-impact chronic pain: a framework for biopsychosocial pain research. Pain 2023;164:2143–2147. doi:10.1097/j.pain.0000000000002951.

[12] Flake JK, Fried EI. Measurement Schmeasurement: Questionable Measurement Practices and How to Avoid Them. Adv Methods Pract Psychol Sci 2020;3:456–465.

[13] Gatchel RJ, Peng YB, Peters ML, Fuchs PN, Turk DC. The biopsychosocial approach to chronic pain: Scientific advances and future directions. Psychol Bull 2007;133:581–624. doi:10.1037/0033-2909.133.4.581.

[14] Giusti EM, Lacerenza M, Gabrielli S, Manzoni GM, Manna C, D'Amario F, Marcacci M, Castelnuovo G. Psychological factors and trajectories of post-surgical pain: A longitudinal prospective study. Pain Practice 2022;22:159–170.

[15] Giusti EM, Manna C, Varallo G, Cattivelli R, Manzoni GM, Gabrielli S, D'amario F, Lacerenza M, Castelnuovo G. The predictive role of executive functions and psychological factors on chronic pain after orthopaedic surgery: A longitudinal cohort study. Brain Sci 2020;10:1–10.

[16] Glymour MM, Weuve J, Chen JT, Glymour MM, Chen JT, Weuve J. Methodological Challenges in Causal Research on Racial and Ethnic Patterns of Cognitive Trajectories: Measurement, Selection, and Bias. Neuropsychol Rev 2008;18:194–213.

[17] Greenland S, Morgenstern H. Confounding in Health Research. Annu Rev Public Health 2001;22:189–212.

[18] Hernán MA. The C-word: Scientific euphemisms do not improve causal inference from observational data. Am J Public Health 2018;108:616–619.

[19] Hernan MA. The C-word: Scientific euphemisms do not improve causal inference from observational data. Am J Public Health 2018;108:616–619. Available: doi:10.2105/AJPH.2018.304337. Accessed 4 May 2023.

[20] Hernán MA, Cole SR. Invited commentary: Causal diagrams and measurement bias. Am J Epidemiol 2009;170:959–962.

[21] Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. Epidemiology 2004;15:615–625.

[22] Hernán MA, Hsu J, Healy B. A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks. CHANCE 2019;32:42–49.

[23] Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol 2016;183:758–764.

[24] Hernán MA, Wang W, Leaf DE. Target Trial Emulation A Framework for Causal Inference From Observational Data. JAMA 2022;328:2446–2447.

[25] Hickling LM, Allani S, Cella M, Scott W. A systematic review with meta-analyses of the association between stigma and chronic pain outcomes. Pain 2024;165:1689–1701.

[26] Hines O, Dukes O, Diaz-Ordaz K, Vansteelandt S. Demystifying Statistical Learning Based on Efficient Influence Functions. American Statistician 2022;76:292–304.

[27] Hünermund P, Louw B, Caspi I. Double machine learning and automated confounder selection: A cautionary tale. J Causal Inference 2023;11:20220078.

[28] Jenkins L, Chang WJ, Buscemi V, Cunningham C, Cashin A, Mcauley JH, Liston M, Schabrun SM. Is there a causal relationship between acute stage sensorimotor cortex activity and the development of chronic low back pain? a protocol and statistical analysis plan. BMJ Open 2019;9:e035792.

[29] Lu H, Cole SR, Howe CJ, Westreich D. Toward a Clearer Definition of Selection Bias When Estimating Causal Effects. Epidemiology 2022;33:699–706.

[30] Lyall DM, Cullen B, Allerhand M, Smith DJ, Mackay D, Evans J, Anderson J, Fawns-Ritchie C, McIntosh AM, Deary IJ, Pell JP. Cognitive test scores in UK biobank: Data reduction in 480,416 participants and longitudinal stability in 20,346 participants. PLoS One 2016;11:e0154222.

[31] McGowan LD. tipr: An R package for sensitivity analyses for unmeasured confounders. J Open Source Softw 2022;7:4495.

[32] Miaskowski C, Blyth F, Nicosia F, Haan M, Keefe F, Smith A, Ritchie C. A biopsychosocial model of chronic pain for older adults. Pain Medicine (United States) 2020;21:1793–1805.

[33] MIRO. Miro online whiteboard (no version provided). 2024.

[34] Ng WQ, Hartanto A. The effect of executive function on the development of chronic pain: A prospective longitudinal study. Soc Sci Med 2022;314:115478.

[35] Noe-Steinmüller N, Scherbakov D, Zhuravlyova A, Wager TD, Goldstein P, Tesarz J. Defining suffering in pain. A systematic review on pain-related suffering using natural language processing. Pain 2024;165:1434–1449.

[36] Pearl J. An introduction to causal inference. International Journal of Biostatistics 2010;6:1–59.

[37] Pearl J, Mackenzie D. The Book of Why: The New Science of Cause and Effect. 1st ed. USA: Basic Books, Inc., 2018.

[38] Peters GY, Crutzen R, Peters GY, Crutzen R. Knowing What We're Talking About : Facilitating Decentralized , Unequivocal Publication of and Reference to Psychological Construct Definitions and Instructions. Meta-Psychology 2024;8:MP.20223638.

[39] Poppe L, Steen J, Loh WW, Crombez G, De Block F, Jacobs N, Tennant PWG, Van Cauwenberg J, De Paepe AL. How to develop causal directed acyclic graphs for observational health research: a scoping review. Health Psychol Rev 2024:1–21.

[40] Rodrigues D, Kreif N, Lawrence-Jones A, Barahona M, Mayer E. Reflection on modern methods: Constructing directed acyclic graphs (DAGs) with domain experts for health services research. Int J Epidemiol 2022;51:1339–1348.

[41] Rohrer JM. Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data. Adv Methods Pract Psychol Sci 2018;1:27–42.

[42] Salonsalmi A, Pietiläinen O, Lahelma E, Rahkonen O, Lallukka T. Contributions of childhood adversities to chronic pain among mid-life employees. Scand J Public Health 2022;50:333–339.

[43] Sánchez-Cubillo I, Periáñez JA, Adrover-Roig D, Rodríguez-Sánchez JM, Ríos-Lago M, Tirapu J, Barceló F. Construct validity of the Trail Making Test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. Journal of the International Neuropsychological Society 2009;15:438–450.

[44] Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, Altman DG, Moher D, Barber R, Denegri S, Entwistle A, Littlejohns P, Morris C, Suleman R, Thomas V, Tysall C. GRIPP2 reporting checklists: Tools to improve reporting of patient and public involvement in research. Res Involv Engagem 2017;3:j3453.

[45] Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, Harrison WJ, Keeble C, Ranker LR, Textor J, Tomova GD, Gilthorpe MS, Ellison GTH. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. Int J Epidemiol 2021;50:620–632.

[46] Traeger AC, Hübscher M, Henschke N, Williams CM, Maher CG, Moseley GL, Lee H, McAuley JH. Emotional distress drives health services overuse in patients with acute low back pain: a longitudinal observational study. European Spine Journal 2016;25:2767–2773.

[47] VanderWeele TJ. Principles of confounder selection. Eur J Epidemiol 2019;34:211–219.

[48] Vila MR, Todorovic MS, Tang C, Fisher M, Steinberg A, Field B, Bottros MM, Avidan MS, Haroutounian S. Cognitive flexibility and persistent post-surgical pain: the FLEXCAPP prospective observational study. Br J Anaesth 2020;124:614–622.

[49] Westreich D, Edwards JK, Van Smeden M. Comment on Williamson et al. (OpenSAFELY): The Table 2 Fallacy in a Study of COVID-19 Mortality Risk Factors. Epidemiology 2020:E1–E2.

[50] Westreich D, Greenland S. The table 2 fallacy: Presenting and interpreting confounder and modifier coefficients. Am J Epidemiol 2013;177:292–298.

[51] Williams A, Lee H, Kamper SJ, O'Brien KM, Wiggers J, Wolfenden L, Yoong SL, Hodder RK, Robson EK, Haskins R, McAuley JH, Williams CM. Causal mechanisms of a healthy lifestyle intervention for patients with musculoskeletal pain who are overweight or obese. Clin Rehabil 2019;33:1088–1097.

[52] Wyss R, Yanover C, El-Hay T, Bennett D, Platt RW, Zullo AR, Sari G, Wen X, Ye Y, Yuan H, Gokhale M, Patorno E, Lin KJ. Machine learning for improving high-dimensional proxy confounder adjustment in healthcare database studies: An overview of the current literature. Pharmacoepidemiol Drug Saf 2022;31:932–943.

8 Figures

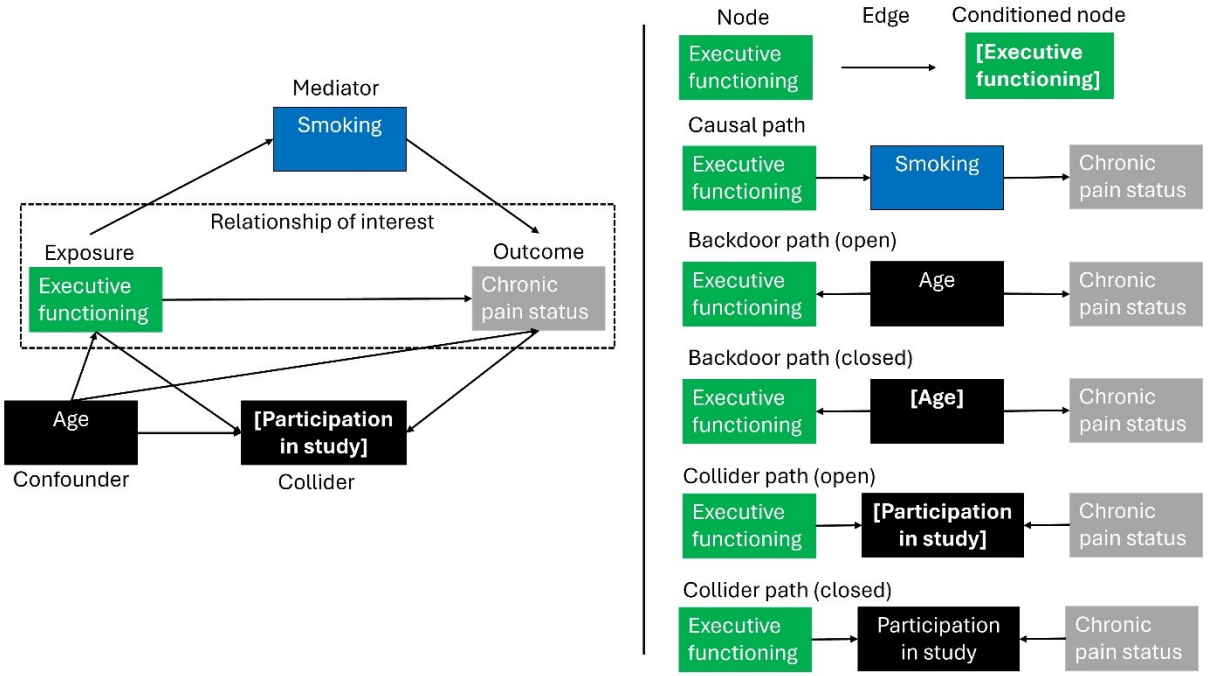

Figure 1. Illustration of the main components of a DAG and the most common types of paths for a hypothetical cross-sectional study in which we want to investigate the influence of executive functioning on chronic pain status (adapted from Tennant et al., 2021). This DAG has been visually arranged so that all edges flow from left to right. To investigate the average causal effect of executive functioning on chronic pain status (relationship of interest) one would have to close the backdoor path via age by controlling for this variable (e.g. by adding it in your regression model). Smoking on the other hand is a mediator and when interested in the total effect (or average causal effect) one should not control for this variable. Since we can only observe data from people who participated in this hypothetical study and participation in the study is, according to this DAG, influenced by both executive functioning and chronic pain status, there is a non-causal path from executive functioning over participation in study (collider) to chronic pain status that remains open and cannot be blocked. This

DAG thus shows that under these assumptions, the causal effect of executive functioning on chronic pain status is not identified.

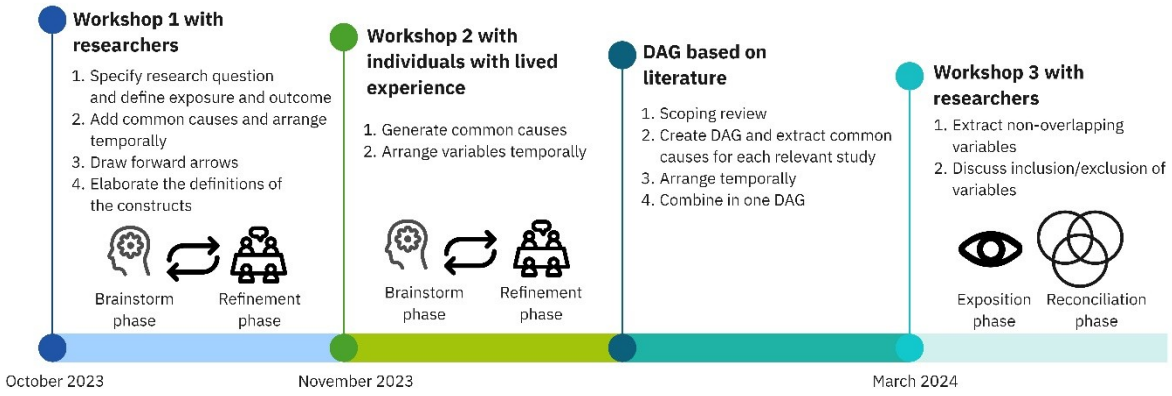

Figure 2. An overview of the different phases of the study.

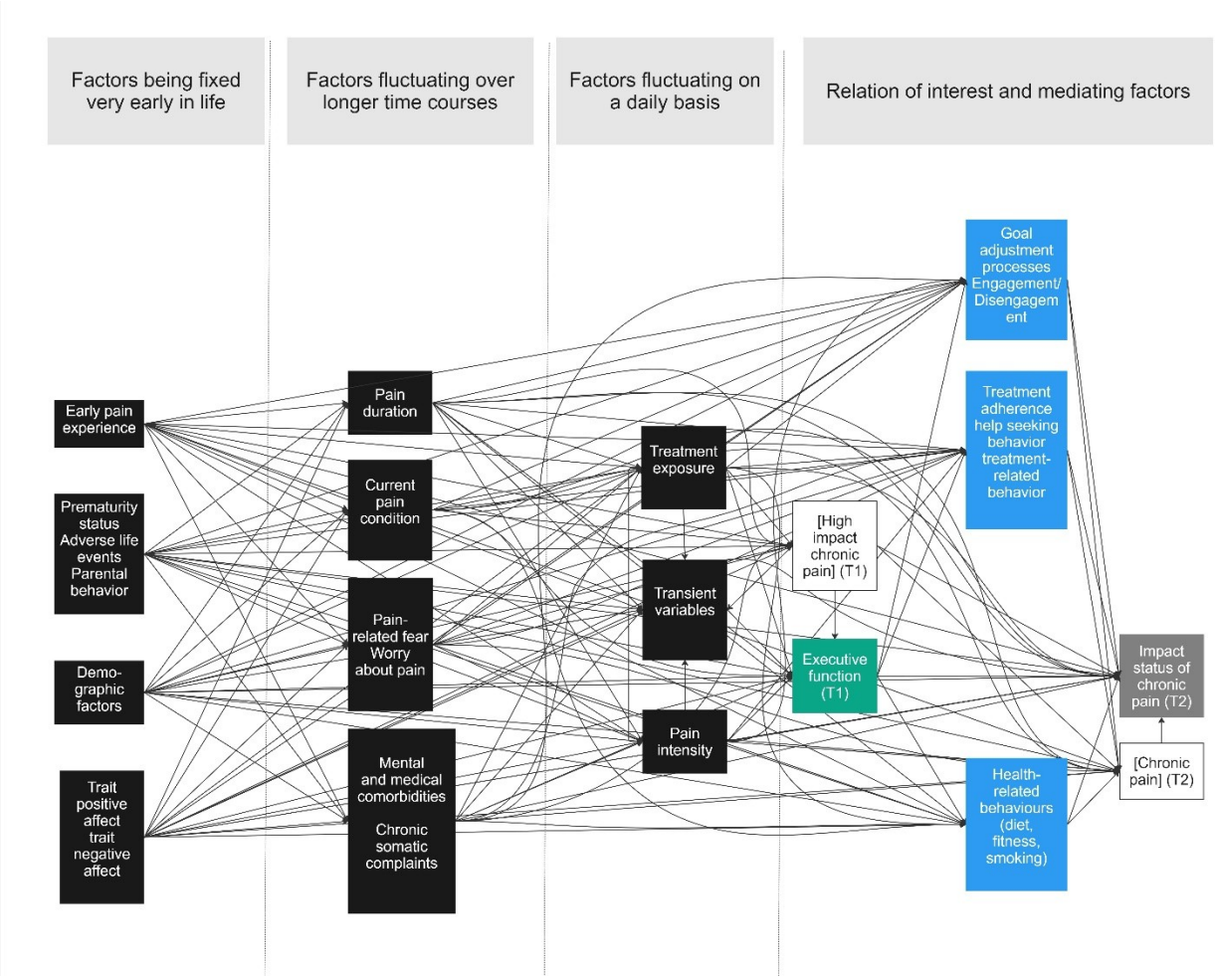

Figure 3. Final DAG for the researcher group at the end of step 3. Green box = exposure; grey box = outcome; blue box = mediator; black box = common cause; white box in square brackets: variables that are fixed at a given value (i.e.

only people with high impact chronic pain were included at T1 and only people with chronic pain at T2).

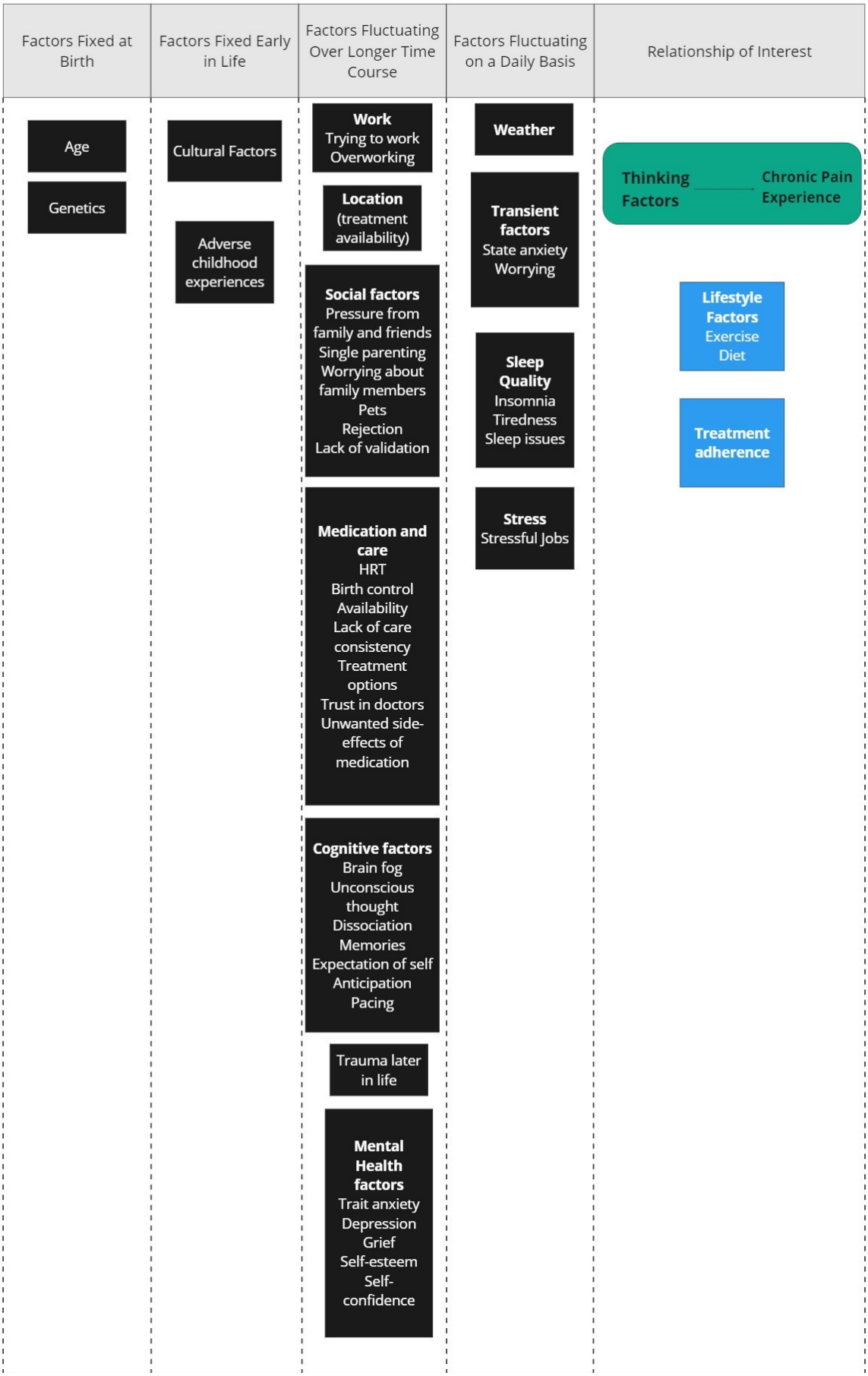
Factors Fixed at Birth
Factors Fixed Early in Life
Factors Fluctuating Over Longer Time Course
Factors Fluctuating on a Daily Basis
Relationship of Interest
Age
Genetics
Cultural Factors
Adverse childhood experiences
Work
Trying to work
Overworking
Location
(treatment availability)
Social factors
Pressure from family and friends
Single parenting
Worrying about family members
Pets
Rejection
Lack of validation
Medication and care
HRT
Birth control
Availability
Lack of care consistency
Treatment options
Trust in doctors
Unwanted side-effects of medication
Cognitive factors
Brain fog
Unconscious thought
Dissociation
Memories
Expectation of self
Anticipation
Pacing
Trauma later in life
Mental Health factors
Trait anxiety
Depression
Grief
Self-esteem
Self-confidence
Weather
Transient factors
State anxiety
Worrying
Sleep Quality
Insomnia
Tiredness
Sleep issues
Stress
Stressful Jobs
Thinking Factors
Chronic Pain Experience
Lifestyle Factors
Exercise
Diet
Treatment adherence

Figure 4. Overview of all the common causes and potential mediators that were identified for the research question of interest by the ILE group. Common causes were put in temporal order by the researchers. Green box = exposure and outcome; blue box = mediator; black box = common cause.

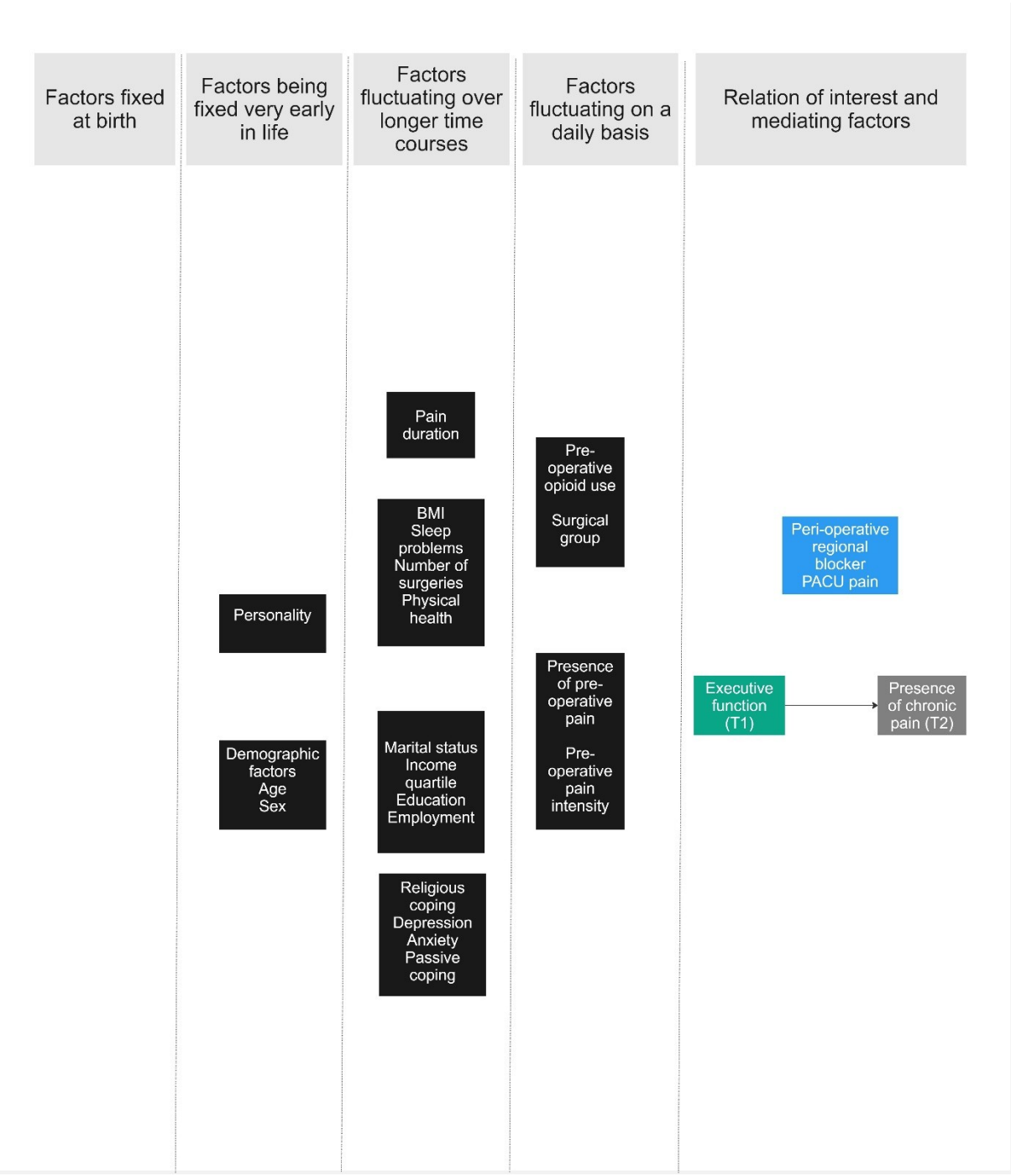

Figure 5. Overview of all the common causes and potential mediators that are potentially relevant for the research question in the literature. Common causes were put in temporal order by the researchers. Green box = exposure; grey box = outcome; blue box = mediator; black box = common cause.

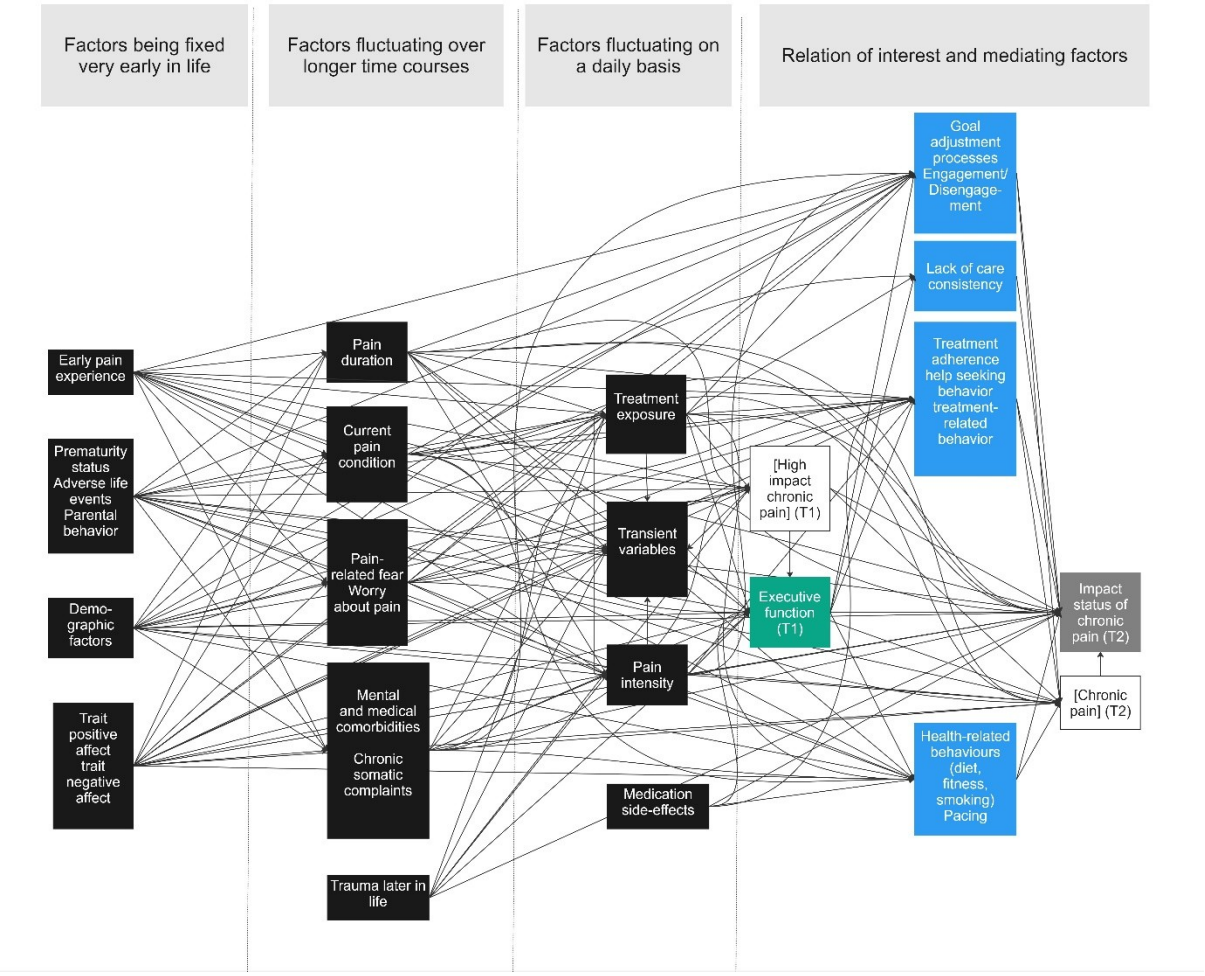

Figure 6. Final DAG after the exposition and reconciliation phase. Green box = exposure; grey box = outcome; blue box = mediator; black box = common cause; white box in square brackets: variables that are fixed at a given value (i.e. only people with high impact chronic pain were included at T1 and only people with chronic pain at T2).