Towards a Data Driven Science of Human Experience

Abstract. By creating a scientific based, version controlled taxonomic language and complimenting research platform for the purpose of pragmatically defining discrete properties of subjective perception in real time, researchers may have access to an empirical framework for conducting first-person experience based quantitative research within the fields of micro-phenomenology, psychology, and potentially many others fields.

Introduction

The traditional scientific field of physics (and others including biology and mathematics) has evolved over thousands of years, thanks in part to the systematic concept taxonomies that grew around those disciplines utilizing empirical methods of observation. As meaningful patterns emerged in relation to how physical objects behaved and interacted, distinct terms were created to codify those observations. For example, friction is a word that was created to describe the force of resistance when an object slides across another. The creation of this word was dependent on people observing certain qualities, and wanting to articulate those qualities exactly, created a word to meaningfully distill that particular phenomenon.

Technical concepts themselves are not stable, and are updated as our understanding evolves. Take for example the physics concept of force (**F**). This concept has existed in some form since before the ancient Greeks, but it continued to go through revision as empirical observations gave us better understanding of how friction relates to force. This is but one word in a vast sea of interrelated technical concepts that has brought the field of physics to where it is today. Through this example we can see that it has taken thousands of years of observations to arrive at the level of discrete definitions we have today in the field of physics. Without other operational concepts such as area, volume, torque, and pressure (just to name a few), the field of physics would not have the expressive and practical power it has today.

When it comes to the science of consciousness, and the vocabulary associated with describing the discrete properties of moment to moment subjective experience, there is no scientifically agreed upon de facto vocabulary to describe this in a measurable, repeatable, quantifiable way, nor is there at the time of this writing a public body which explicitly expresses the intention to foster the creation of it.

As the scientific community around consciousness research continues to advance, the need for a set of discrete, **evolving** terms to describe observations that are completely subjective in nature (e.g. only available and describable directly by the actual observer) becomes a necessity if we hope to develop a meaningful, universally agreed upon set of fundamental units and concepts to articulate the dimensions and properties of subjective perception. This paper intends to propose a path forward in directly quantifying and measuring those dimensions and objects of subjective consciousness directly.

In order to effectively follow along with what is being proposed, this paper will utilize a collection of terms that intend to not only refer/point to very specific abstract ideas, but also to point to, reify, and conceptualize specific qualities of immediate, direct, conscious experience. Below is a list of these terms as it relates to the content of this paper. Importantly, as consensus grows, these terms are subject to change in meaning and therefore ways of managing this varying dynamic is built directly into the paradigm this paper will be proposing.

Sense Phenomena: X

This umbrella term is used to refer to all/any type of subjective perception that can possibly be defined/experienced, be it body sensations, emotions, thoughts, etc. It should be made clear at the outset that although certain categories of **sense phenomena** will likely be created to most effectively define sensory experience at the lowest level of abstraction, this paper do not claim that any category is ontologically "real", only that it is practical to create such categories for the purposes of investigating some particular perception in detail. The symbolic representation for generic **sense phenomena** is **X**.

Sense Phenomena Class

A generic term used to refer to a specifically defined type/category of **sense phenomena**, with its own unique set of properties**.** Within the paradigm proposed, researchers can create classes of **sense phenomena** in order to observe/measure a particular aspect/aspects of **conscious space**.

This term is borrowed from the computer programming concept "Class", and many of the ideas associated within programming apply here too, most specifically the concept of inheritance. This means that there can technically be **sense phenomena sub-classes,** or types of **sense phenomena** that inherit the meaning of other **sense phenomena classes**, along with their own unique properties.

Conscious Space

This term refers directly to the totality of the perceivable field in which all **sense phenomena** appear. To give a direct example, the very **sense phenomena** associated with perceiving these words on this document in this very moment is arising within **conscious space**. It is that which contains the totality of all dimensions of what is consciously experienced in this moment. It would appear initially grammatically more concise to assign an owner to **conscious space**, such as to say "your" **conscious space**, however, given the paradigm that is being purpose, any **sense phenomena** associated with a felt sense of "self" is ultimately contained within **conscious space** itself. For this reason, possessive terms will not be used while referencing **conscious space** in this paper.

Observer

This term is used to refer to an individual which is aware of **sense phenomena** occurring. **Observers** observe **conscious space** and serve as the primary means of reporting within this paradigm.

Verbal Thought: VT

An arbitrary **sense phenomena class** which is focused on measuring the manifestation of language content within **conscious space**. Examples of verbal thought include the sensory experience of thought, reading out loud in your head, or automatic thoughts which arise outside volitional control. Verbal Thoughts can arise with and without a sense of intentional will. Cognitive behavioral therapy has a particular interest in thoughts that arise without intentional will, referred to as automatic thoughts. The symbolic representation of **verbal thought** is **VT**.

Manifested Sense Phenomena: (M)

This term is a generalized word used to refer to **distinct temporal instances** of any particular class of **sense phenomena** occurring. The symbolic representation of **manifested sense phenomena** is **(M)**. **(M)** will ultimately become the primary source of research data within this proposed paradigm as it is directly observable and reportable.
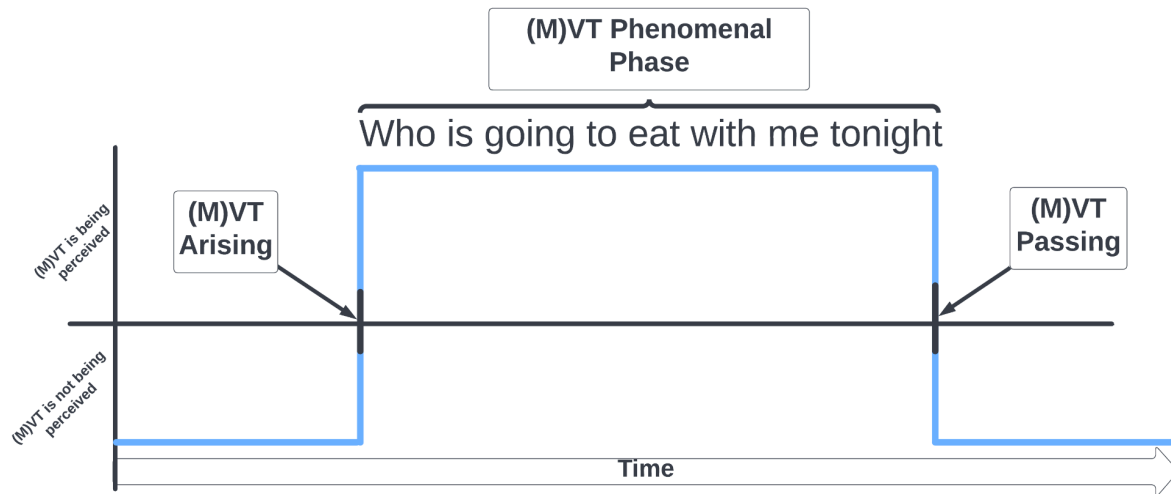
Manifested Verbal Thought: (M)VT

Figure 1:
This figure depicts a single **(M)VT** arising and passing within **conscious space.**

(M) Arising

This term points to the temporal moment that a particular class of **(M)** first arises within **conscious space**. It should be noted that for each class of **(M),** the definition of **arising** may be further defined, or not apply at all.

(M) Phenomenal Phase

This refers to the period of time between the **arising** and **passing** of **(M)X.** In some sense, the **phenomenal phase** is completely arbitrary, for example, if observing a **verbal thought**, there are technically spaces of phenomenal inactivity between the perceived individual words associated with a **verbal thought**. Given this observation, for each class of **(M)**, the **phenomenal phase** will be more discreetly defined.

(M) Passing

This term points to the temporal moment that a particular class of **(M)** disappears within **conscious space**. It should be noted that for each class of **(M),** the definition of **passing** may be further defined

(M) Value

This term refers to the actual content of the **(M)X.** The **value** of the **(M)VT** within figure 1 would be the english words "Who is going to eat with me". All instances of a particular **(M)** that have the exact same **value** would be considered the "same".

Value Format

This term refers to the data format and constraints that the **(M) value** is encoded in**.** For example, for **(M)VT**, the **value format** is text (all languages). A subtle but important facet to point out is that the **value format** of a particular **(M)** is a transposition of the direct experience of that particular **(M)**. What this essentially means is that the direct experience (which is **only** available to the **observer)** is converted into some meaningful/pragmatic representation of the original direct experience.

Format Value Transposition Rule: FVTR

This term refers to the rules that should be followed when transposing the direct experience of **(M)** into a **value format**. These rules are intended to make decisions on how direct experience should be encoded into the **(M) value** for the sake of making the data uniform and disambiguous. Take for example the following **(M)VT** content:

- I'm going to be twenty two tomorrow!
- I'm going to be twenty-two tomorrow!
- I'm going to be 22 tomorrow?

When a **VT** arises, there may be no perceivable punctuation other than the intonation that is directly experienced. Rules within this context are used to decide how these three different **(M)VT**s should be classified. Are all three the same? Are all three different? FVTRs are used to resolve these types of conflicts.

Abstract Sense Phenomena: (A)

This concept serves as a reference point for all instances of a **manifested sense phenomena** that have the same **value.** The symbolic representation of **abstract sense phenomena** is **(A)**. The utility of an **(A)** and the reason for defining them as a concept is because within the graph model that will store data within this paradigm, an **(A)** serves as the means to connect **(M)X**s that contain the same, but temporally separated **values**. The figure below helps to articulate this meaning.
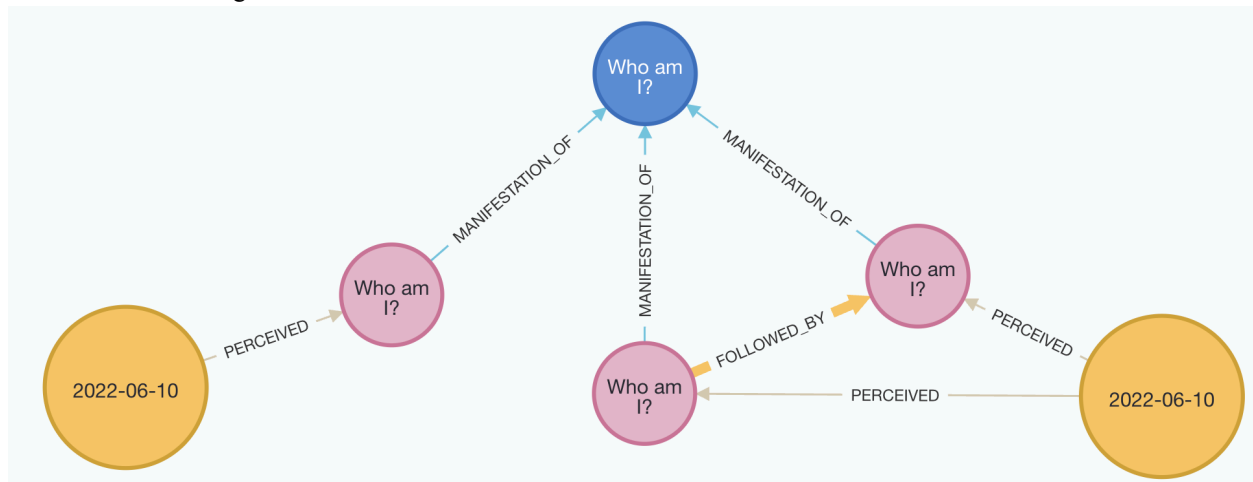

Figure 1:
The blue node is a graph representation of an **Abstract Sense Phenomena.** The three pink nodes are the **(M)VT**s "Who am I?". The yellow nodes represent unique sessions of observation. **Abstract Sense Phenomena** serve to connect separate instances of **manifested sense phenomena.**

An analogy can be made between the field of chemistry. The element hydrogen is an abstract concept, whereas the hydrogen gas atoms that are present within **conscious space** right now are manifestations of that abstraction.

Sense Data

This term is a generalized word used to refer to any **sense phenomena** that is collected by **observers**.

This concludes the most immediately relevant definitions required to understand this paper. Additional terms will be introduced as they become relevant. With that said, this paper proposes several unique but connected ideas which will need to all come together in order for this project to be successful:

1. The development of a taxonomic, version controlled, symbolic language for the description of observed **sense phenomena** and their associated properties, units of measurement, etc. For brevity's sake, this shall be referred to as the **SPSL** (sense phenomena symbolic language).
2. The creation of **sense phenomena** reporting techniques that allow individuals to report into an arbitrary human-computer interface (eg, keyboard, touchscreen, etc) as **observers** of the **sense phenomena** that are occurring in real time within **conscious space,** utilizing the **SPSL**. For brevity, this component shall be referred to as **SR** (sense reporting).
3. The development of a version controlled, property graph model on which the **SPSL** can be mapped, thus allowing representations of sensory experience to be stored within this graph, enabling graph based data analysis. For brevity, this shall be referred to as the **SP-GM** (sense phenomena graph model).
4. The creation of units of measurement for describing **sense data**
5. The creation of research procedures, methods, and protocols which would allow individuals to report **as observers** using **SR** in scientific research
6. The creation of a software platform that enables all of these pieces to come together for the purposes of conducting research into **SR,** maintaining the **SPSL,** and lastly, serving as the means by which the **SPSL** evolves over time.

Sense Phenomena Symbolic Language (SPSL)

The first proposal is the development of a taxonomic, version controlled, symbolic vocabulary for the purpose of quantifying and identifying the discrete properties of direct subjective perception/**sense phenomena** as it arises and passes in real time. The intent of this language is to provide discrete terms which can be used to categorize different **sense phenomena** as they occur, and associate with those specific phenomena their distinct qualities such as duration, hedonic tone, and location within **conscious space**. A comparison would be between the **SPSL** and the Standard Model in particle physics, which describes the elementary particles (categories), as well as their unique properties.

Put broadly, the **SPSL** intends to serve the following goals:
● Be ontologically neutral
● Allowing for the creation of arbitrary **sense phenomena classes** for the purposes of measuring desired dimensions of experience
● Be mappable onto higher order categorizations of subjective perception (such as cognitive behavioral therapy)
● Provide mechanisms for consensus to update terms as the field evolves and meanings become more refined
● Develop a system of symbols to represent concepts in the **SPSL** for the sake of communicating certain ideas exactly
● Serve as a universal language for describing the properties within **conscious space**

To further elaborate on the idea associated with creating arbitrary **sense phenomena classes**, it's important to stress that a major goal of what is being proposed in this paper is to enable researchers to prototype how phenomenal content is defined, for the purpose of measuring it. What is an emotion for example? There are many different meanings behind this word, and by enabling researchers to create arbitrary **sense phenomena classes**, they can explicitly define the boundaries around what subjective experience they're intended to capture, without having to argue over whether or not their definitions are "correct".

For example, let's say that there is a desire to investigate emotion. A **sense phenomena class** could be defined as **Emotion,** with a symbolic representation of **E. E** could be defined as a "felt sense" tone that can occur technically anywhere in **conscious space**. The felt sense of anger, happiness, sadness, and fear could be defined as examples of **E**. Another team of researchers may be more interested in looking at higher order emotions that are relational, such as jealousy, contempt, and envy. This could be defined as **Relational Emotion**, with **RE** as its symbol.

The **SPSL** intends to serve as a means of allowing researchers to define **sense phenomena** in whatever way they believe is practical for what they are striving to measure. The **SPSL** is not limited to simply thought and emotion, any perceivable **sense phenomena** can be defined, it all depends on what the researcher is interested in measuring.

Assuming this approach to understanding subjective perception proves beneficial, we need a way to anticipate changes in our understanding of **sense phenomena** overtime. With that in mind, the **SPSL** will utilize version controlling. This is so that as the community around this research grows, consensus can be used to introduce new concepts and better descriptions, which can thus be added to the **SPSL** in distinct versioned releases. Version controlling will also allow for research data to be explicitly defined within the parameters of the **SPSL** version that is used to collect data.

Sense Reporting (SR)

By creating the **SPSL,** it provides us a language which can be utilized to actually capture **sense phenomena** in near real time, with the latency being the time duration between immediate awareness of the sense phenomena, to the beginning of input to an input device (EG, keyboard, etc). **Sense Reporting (**SR) refers to any techniques that utilize the **SPSL** in order to capture **sense phenomena** in a way that can be digitized and referred to later for analysis. This technique is inspired by the meditation tradition known as vipassana, and more specifically Mahasi Sayadaw's noting technique as well as Shinzen Young's Unified Mindfulness, in which as **sense phenomena** arises, it is classified/categorized. Below is a diagram which visualizes an example **SR** session.
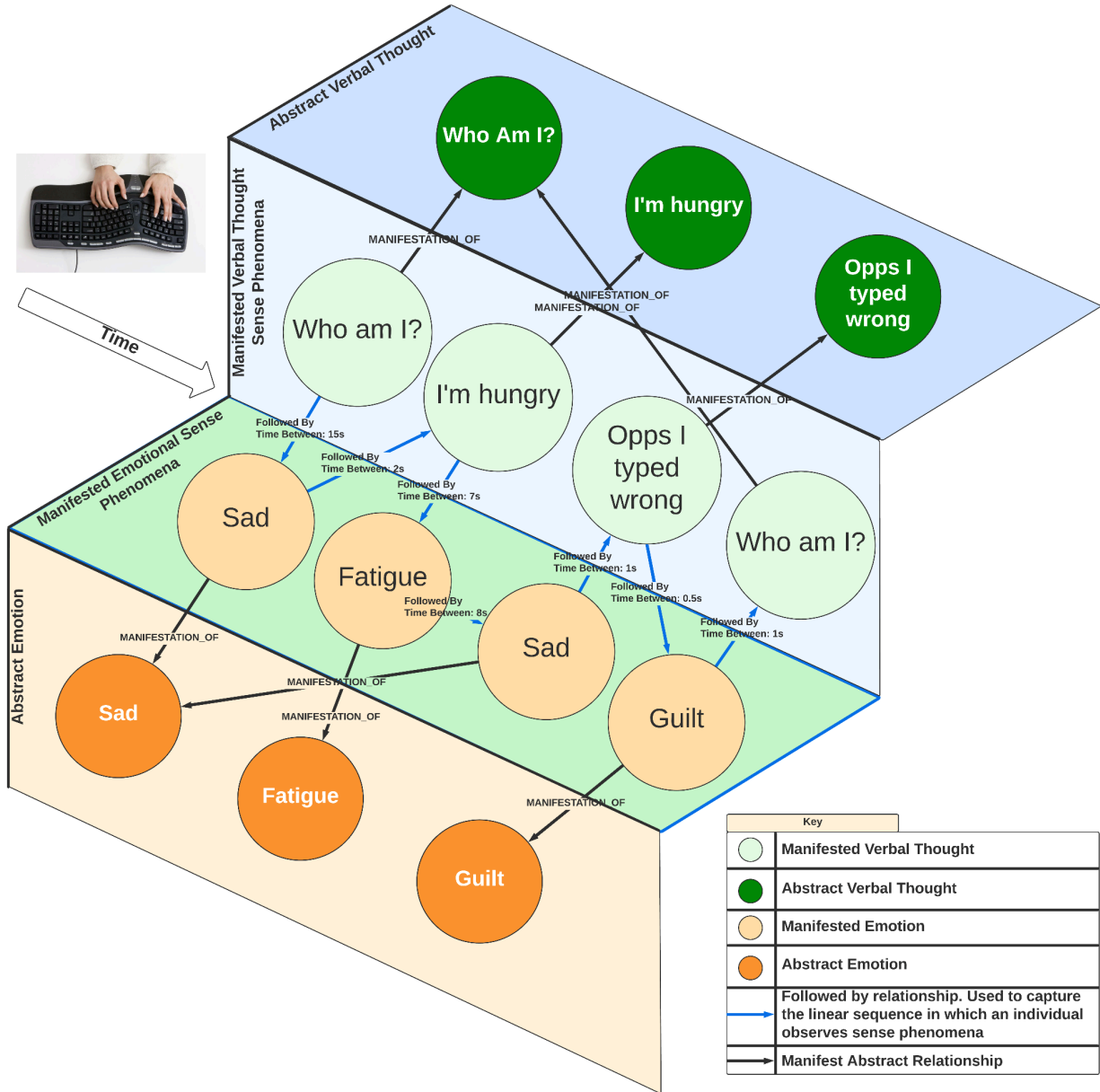
Figure 2:
The above graphic demonstrates an **observer** conducting a **Sense Reporting** session, tracking the flow of **Manifested Verbal Thoughts** and **Manifested Emotions** as they arise. The data is natively stored in a graph for future analysis.

The most basic form of **SR** is **passive thought tracking (PTT).** PTT is a practice in which an **observer** sets a timer (typically 5 minutes, but can be longer or shorter depending on the experiment), and during that period of time, the **observer** witnesses **conscious space**. As **(M)VT** occurs, they are captured via an input device, typically a keyboard. At the same time, various properties of the thought may also be reported, such as the associated hedonic tone.
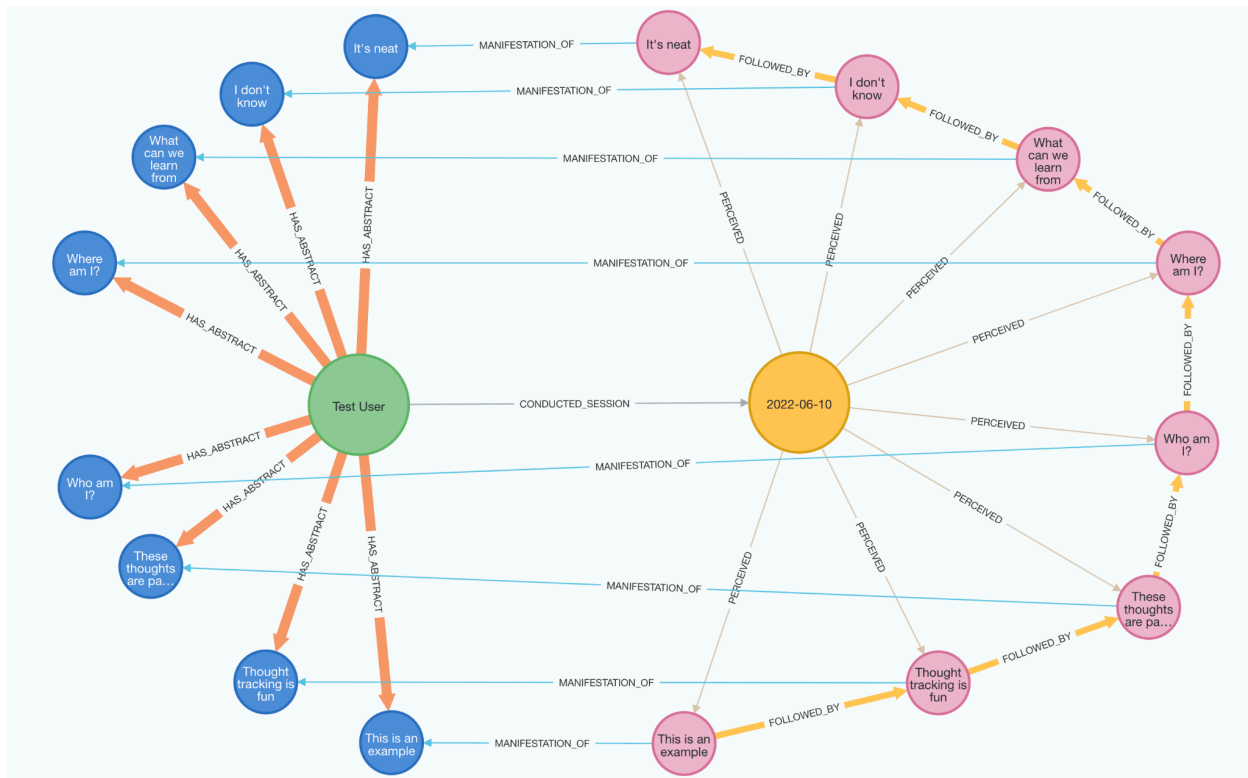
Figure 3:
This figure shows an implementation of the **SP-GM**. The green node refers to the individual **observer (Test User)** tracking the stream of automatic **(M)VT** arising. The yellow node represents an individual instance of a **sense reporting session (SRS).** Pink nodes refer to each unique **(M)VT** that arises during the **SRS**. The blue nodes serve as **(A)VT** to keep track of every single time a particular **(M)VT** arises.

To explain **sense reporting** in a practical context, let's take a real world example of how this could work. Within the world of psychedelic research, a major question is on what type of effects these psychoactive substances have on the individual. As part of a 12 week study in which an individual is given three MDMA sessions in tandem with 12 weeks of therapy, **SR** could be incorporated into the experiment design. The individual would be required to conduct a 5 minute **PTT** session at a specific time during the day. At the end of the 12 week period, this **sense data** can be correlated with the Clinician-Administered PTSD Scale for DSM-V (CAPS-5) that is administered at the beginning and ending of the trial period.

Sense Phenomena Graph Model (SP-GM)

In order to make **sense data** meaningful for analysis, it behooves us to have a standardized model by which this information can be represented computationally. The purpose of the **SP-GM** is to serve as a standardized, version controlled property graph model for structuring **sense data**. This serves several purposes; first, standardizing the way in which **sense data** is digitally structured supports reproducibility of experiments which utilize the **SP-GM**. Additionally, version control enables researchers to be explicit in how the data should be interpreted. Lastly, property graph models allow for very complicated, highly relational queries, as well as enabling graph traversal, which can allow us to learn more about the relationship between **(M)X** such as:

- For a given **(M)X**, what **(M)X** most frequently immediately precedes that **(M)X**? What about the 2nd, 3rd, and on?
- For a given **(M)X**, what **(M)X** what most frequently follows immediately after that **(M)X**? What about the 2nd, 3rd, and on?

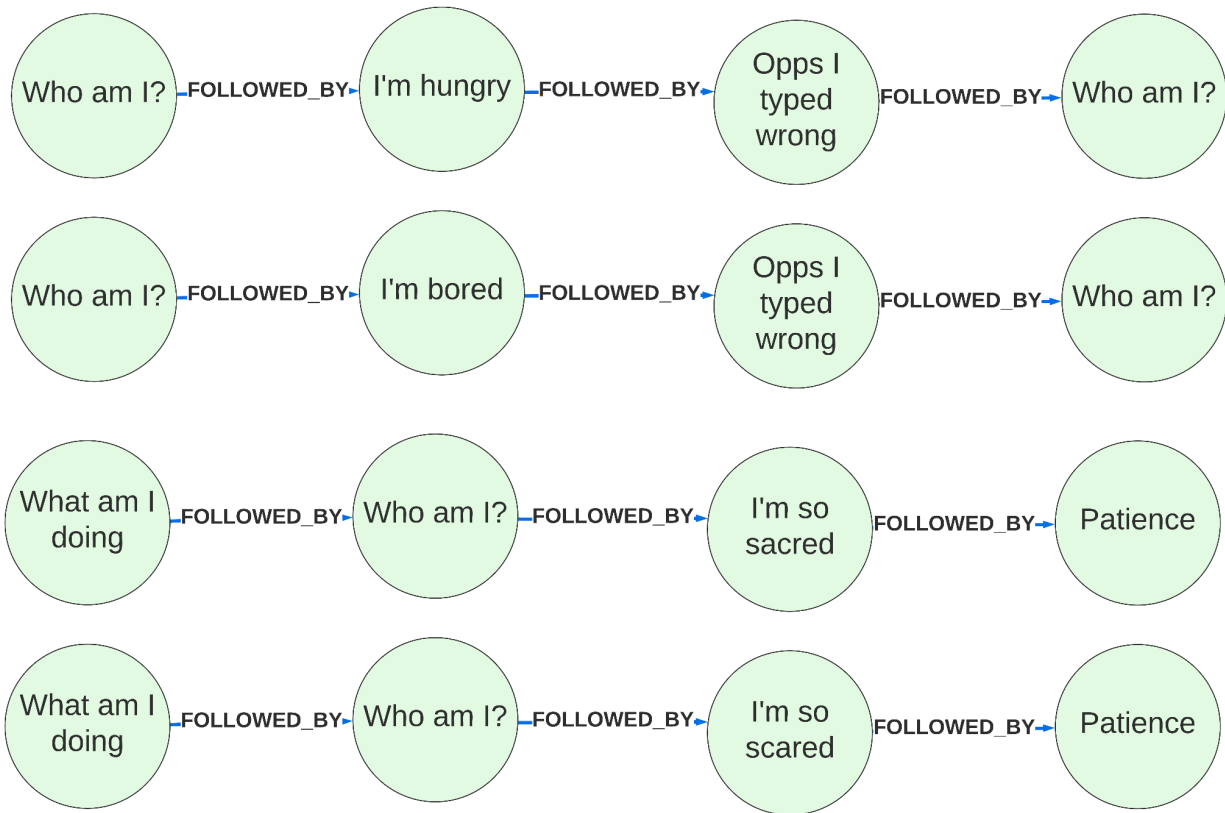An example of traversal can be seen below:

Figure X:
The green node refers to **(M)VT**, with each row representing a unique **reporting session.** The **FOLLOWED_BY** relationship stores the temporal relationship between distinct **(M)VT**. Traversal helps us to identify the relationships and patterns between the sequence of arising **(M)VT,** such as the fact that the most frequently preceding **(M)VT** for "Patience" is "I'm so scared".

Sense Reporting Platform (SRP)

The last major proposal focuses on the development of a software research platform which brings all of these distinct ideas together. This platform would serve as the central hub for the following activities:

- **Sense reporting** experiential design creation for researchers
- Double blind study support
- Data collection interface for individual **sense reporting** sessions
- Biosignal data correlation with **sense reporting**
- Data analysis
- Entry and management interface for **SPSL** vocabulary
  Version control of collective consensus data for adding new terms/update old definitions
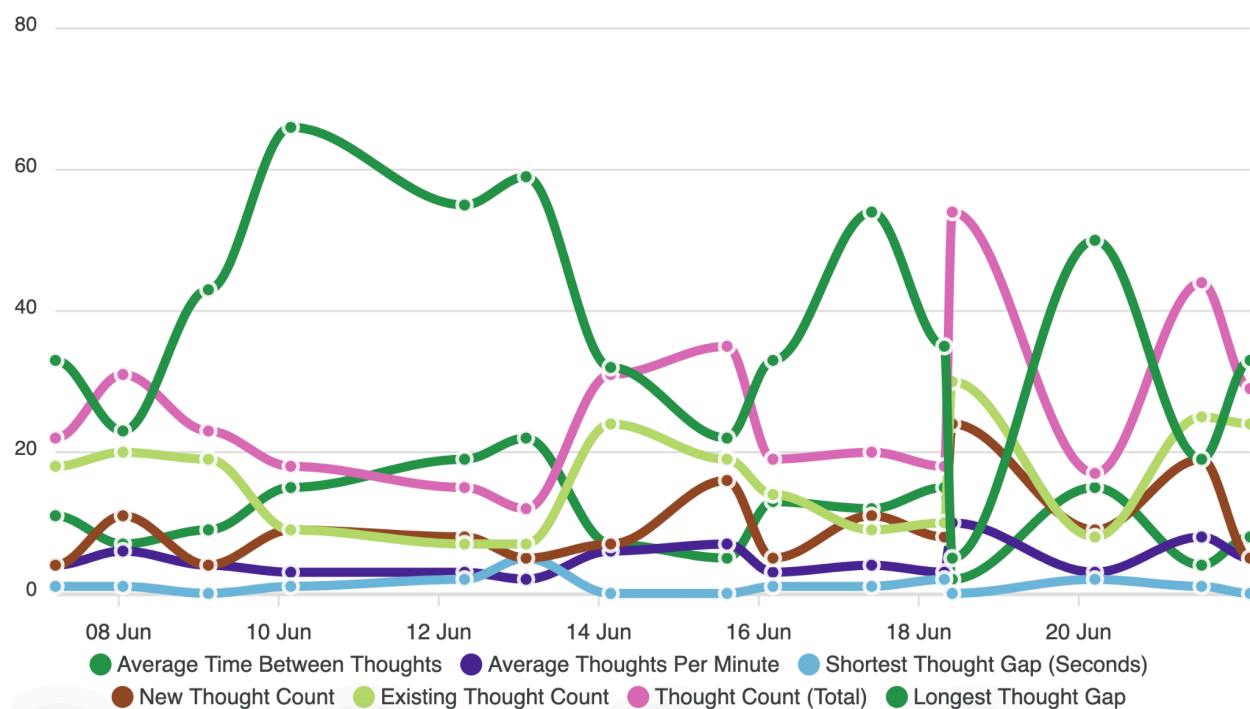- **Centralized, sense data** warehousing

Units of Measurement

Figure 4:
This diagram represents a time series visualization of unique **sense reporting** sessions for **(M)VT**, and their associated data. The **X** axis represents time in days, the **Y** axis represents the amount reported for a particular unit of measurement for a distinct session.

Within individual **sense reporting** sessions, we can create units of measurement to describe certain qualities of the activity of **sense phenomena.** These units of measurement will continue to refine as the field evolves. With that said, the list below (specifically for **(M)VT**) demonstrates a subset of the units of measurement that can be extracted out of **sense reporting.**

Average Manifested Verbal Thoughts per Minute: (M)VT/m
(M)VT/m refers to the average number of **(M)VT** per minute reported during a **SR** session.

Average Time Between (M)VT
Within a **SR** session, the average time gap between reported **(M)VT**.

New (M)VT Percentage
Within a **SR** session, the number of new **(M)VTs** (never reported before) that appear in the session.

Existing (M)VT Percentage
Within a **SR** session, the number of **(M)VTs** reported that already have an existing **(A).**

Applications

The **sense data** produced by **sense reporting** can be applied scientifically in a variety of ways. The list below shows a few applications in which this protocol could yield meaningful data.

Psychology Comparative Studies

With access to subjective states, **SR** provides the ability to directly quantify the content of **conscious space.** Comparative studies could be conducted with **SR** between individuals with mental disorders (such as depression, OCD, PTSD, schizophrenia) and those that consider themselves to be clinically healthy. What would the differences be between these two groups? What would the similarities be?

Efficacy Studies / Clinical Trials

Any research which intends to measure the efficacy of a particular health intervention can incorporate **sense reporting** in a number of ways. For example, The MAPS PTSD clinical trial consists of a 12 week treatment period in which the patient is given 3 sessions with MDMA in order to treat PTSD. If the protocol included daily 5 minute **sense reporting** sessions during the 12 week period, this data may serve to develop a deeper understanding of the relationships between improvements, and changes in the content of **sense phenomena** over that period of time. Questions that could be asked include:

- What is the relationship between symptom improvement and changes in an **observer's** reported **sense data** throughout the period of the trial?
- What specific changes in **sense data** (if any) correlate with individuals that reported improvements at the end of the trial?

Measuring the effects of meditation techniques

Do meditation techniques have specific effects on the state of **conscious space**? Experiments can be designed in which an **observer** measures their thought activity prior to the meditation, and then immediately after the meditation. Could it be proven that certain meditation techniques have an impact on the **(M)VT/m?**

Researching the relationship between sense phenomena and DSM diagnostic criteria

What is the relationship between an individual's reported **sense data** and the likelihood that they meet the DSM standard for severe depression?

Physiological correlates associated with the flow of sense phenomena

Does **(M)VT/m** have a correlation with specific physiological activity? Heart rate, EGG, skin conductance, etc.

Research focused on individuals with extensive meditative experience

What does the activity of **sense phenomena** look like in someone that has meditated daily for 20 years compared with someone who has meditated 1 year? Do long term meditators have different **M)VT/m** values? Do experienced meditators have lower **(M)VT** activity compared to non-meditators? Do they perceive more?

Machine Learning Training Data

There already exists a wide variety of biosignals that can be collected on individuals, such as EEG, FMRI, heart rate, and electrodermal activity. It is theoretically possible that **sense data** in tandem with biosignal data could be tagged and treated as training data for creating machine learning models that could be used to predict properties of **(M)X** based on biosignal data.

Basic Research into Sense Phenomena

Many of these applications have applied to already existing fields of research, but there are those that may be interested in this field for the sake of exploring and expanding methods, techniques, and concepts for investigating **sense phenomena** directly.

Ultrasonic Neuromodulation Research

The SEMA lab headed by Dr. Sanguinetti and Shinzen Young currently focuses on the effects that ultrasonic neuromodulation can have on meditative practice. By applying **sense reporting**, it is possible that the subjective effects of this neuromodulation may be more directly quantifiable.

CBT, Internal Family Systems, Ideal Partner Protocol, and other psychotherapies

The field of psychology contains many different forms of psychotherapy. With the creation of the **SPSL**, these different modalities potentially serve to benefit from a deeper understanding of the raw **sense data** that is associated with their respective disciplines. Combining these therapies with **sense reporting** could potentially serve to deepen the therapeutic relationship, and research could be conducted to determine the effects of these therapies as it relates to emotion state, thoughts, etc.

Conclusion

The goal of this paper is to serve as a starting point for discussion and to lay out a formal system by which the properties within **conscious space** can be researched, quantified, and reported on. Thus, the intention has been to express that this is ultimately an **evolving** paradigm, with mechanisms in place to help account for changes in our understanding as the discipline matures overtime (eg, version controlling and consensus voting).

References

BELOW THIS POINT IS JUST NOTES AND UNFINISHED IDEAS, PLEASE IGNORE

This paper was intended to serve as a starting point for discussion, and to solicit the attention of those that may have an interest in utilizing this paradigm for research purposes. Many steps and individuals are needed to materialize this vision, but from a technical perspective, there is nothing inhibiting the creation of this paradigm, other than a lack of curiosity in what types of questions and data may come out of this approach.

Extra Text with no home yet

The goal of this outline is to evangelize the development of a neutral, atomic language to quantify subjective phenomena that enables it to map onto higher order models of subjective experience, such as Cognitive Behavioral Therapy, Schema Therapy, and Internal Family Systems (IFS).

- Abstract
- Introduction
    - Describe the utility and why of how this is relevant (Mental illness increase over the years?)
    - Present the main components of subjective perception tracking
- Subjective Perception Tracking Method

- - - Need to articulate how each session has an associated "Template" which defines the constraints for the particular session (what is tracked, how its tracked, etc)
  - Subjective Perception Language
    - Explain how the SPL is intended to help us quantify unique dimensions of conscious perception,
    - Visualize exactly how version controlling will work (Show how previous versions will be maintained
  - Graph Model
    - Explain how using graph theory, we can explore unique relationships we couldn't otherwise describe
  - Applications
    - Outcome Studies
  - Create a visualization of the relationship between:
    - **SPSL**
    - Sense Reporting
    - Graph Model
    - Software Platform

Existing Techniques for studying subjective experiences:
1. Experience Sampling Method (found of Flow state)
2. Descriptive phenomenological method (Amedeo Giorgi)

References:

- Awesome Paper by Terje Sparby about the future of contemplative science
  https://www.researchgate.net/publication/318207094_The_Nature_of_Contemplative_Science_and_Some_Prospects_for_its_Future_Development
https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01980/full

State of Self Reporting: https://onlinelibrary.wiley.com/doi/full/10.1111/sjop.12671

https://elight.springeropen.com/articles/10.1186/s43593-022-00019-x

EEG Based Speech Recognition: https://core.ac.uk/download/pdf/197555837.pdf

Philosophy of Micro Phenom: https://philpapers.org/rec/PETOTV

Additional applications that are incomplete:

Emergent Phenomenology Research (Incomplete)

Philosophy of Sense Phenomena (Incomplete)

Self Knowledge (Incomplete)

Breathwork Research (Incomplete)

Breathwork refers to the whole collection of breathing modulation techniques. Experimental designs can be set up with control groups where volunteers conduct sense reporting sessions prior to and post breathwork sessions. This

Experimental Memory Recall Research (Incomplete)

Technically speaking, any training program, be it academic courses, micro learning, all attempt to store information in an individual's long term memory. **Passive Thought Tracking** could be combined with guided recall queues to see what **(M)VT**

Drug Effect Phenomenology (Incomplete)

With the doubt cast on the serotonin theory of depression (Moncrieff, 2022), coupled with the fact that antidepressants are still better than placebo, the question emerges of what are these types of drugs doing that facilitates their antidepressant effect? **Sense data** can be used in combination with medication, and

Psychology/Education Intersection (Incomplete)

What data can we correlate between mental state activity and ability to retain information? Can you mental activity or "distractavility" be quantified in such a way that you can have people do learning tasks where their distracitibility is known before

Introducing Phenomenology into the Paper