

# Protocole expérimental sur la perception humaine des essaims

Aymeric Hénard

July 17, 2023

## Abstract

Ce document détaille le protocole d'expérimentation permettant d'étudier la perception humaine des essaims, et plus précisément comment les humains perçoivent les fractures pouvant se produire dans l'essaim.

## 1 Vocabulaire

- **Fracture de l'essaim** : Perte de la connexion sociale (communication et coordination, rupture de l'agrégat) entre un agent ou groupe d'agents et l'essaim principal.
- **Groupe F** : Agent ou groupe d'agents quittant l'essaim lors d'une fracture. Plusieurs groupe F peuvent exister simultanément.
- **Clip** : Une vidéo montrant un essaim évoluant dans son environnement (plus précisément l'affichage des log d'une simulation d'essaim).

## 2 Contexte théorique

Dans le contexte de essaim de robots, les fractures font références à des ruptures de communication et coordination parmi les robots, pouvant réduire l'efficacité de l'essaim voire même l'empêcher d'accomplir l'objectif qui lui a été confié. Une fracture peut apparaître à cause de diverses raisons, comme des changements de l'environnement, des pannes parmi les robots, des objectifs différents ou bien un mauvais contrôle de la part de l'opérateur. Dans le contexte du contrôle d'un essaim par un opérateur, on souhaite l'aider à contrôler l'essaim sans provoquer de fracture, néfaste quand celle-ci n'est pas voulue.

## 3 Objectif de l'expérimentation

### 3.1 Objectifs

- Vérifier si un utilisateur a besoin d'informations supplémentaires pour percevoir et prévoir les fractures des essaims, dans le but de maintenir l'intégrité de l'essaim en empêchant les fractures.
- Identifier les critères rendant difficile ou plus facile la perception des fractures, si ceux-ci existent. Cela peut faire l'objet d'une expérimentation à part entière. En conséquence, l'identification de ces paramètres, notamment provenant du questionnaire et de la discussion avec les participants, sera l'ouverture vers une analyse plus poussée, probablement pas réalisable dans le cadre de cette thèse. Cependant, deux critères seront évalués dans l'expérimentation : la distance de fracture et la vitesse.

### 3.2 Questions de recherche

L'expérimentation présentée dans ce document a pour objectif d'étudier la perception humaine des essaims en cas de fracture, et plus précisément de répondre aux questions suivantes :

- Un humain perçoit-il les fractures de l'essaim?

- Un humain perçoit-il les fractures avant que celles-ci n'apparaissent?
- Le comportement de l'essaim a-t-il un impact sur la perception de la fracture par un humain?
- La distance entre les agents a-t-elle un impact sur la perception de la fracture par un humain?
- La vitesse d'éloignement du groupe responsable de la fracture a-t-elle un impact sur la perception de la fracture par un humain?

### 3.3 Hypothèses

Cette section regroupe les différentes hypothèses de recherche que l'on souhaite vérifier grâce à cette expérimentation. Voici les différentes hypothèses envisagées :

1. Les humains perçoivent les fractures des essaims.
2. Les humains ne perçoivent pas les fractures avant que celles-ci n'apparaissent.
3. Les humains ne perçoivent pas le moment/seuil exact de la fracture de l'essaim.
4. Les humains perçoivent mieux les fractures quand la distance inter groupes est supérieure à la distance intra groupe.
5. Les humains prévoient mieux les fractures quand la distance inter groupes est supérieure à la distance intra groupe au moment de la fracture.
6. Les humains perçoivent mieux les fractures quand le groupe responsable de la fracture s'éloigne rapidement du groupe principal.
7. Les humains prévoient mieux les fractures quand le groupe responsable de la fracture s'éloigne rapidement du groupe principal.
8. Certaines catégories provoquent chez un humain un jugement de prédiction différent des autres catégories (peut importe la réponse).
9. Certaines catégories sont plus difficiles à percevoir pour un humain.
10. Certaines catégories sont plus facile à prédire pour un humain.

## 4 Méthode / Protocole (durée)

### 4.1 Population

- 1 groupe de personnes lambdas
- Tranche d'âge visée : 18 ans et +
- Nombre : au moins 30 personnes

### 4.2 Matériel

- Un ordinateur portable "Alienware m15 R5 Ryzen" avec un écran de 1920 x 1080, 165Hz

### 4.3 Software

- Implémentation des boids de Reynolds (avec des forces virtuelles).
  - Règle d'attraction.
  - Règle de répulsion.
  - Règle d'alignement.
  - Règle de déplacement aléatoire.

- Système accélération - vitesse - position.
  - Règle de friction.
  - Règle d’anti-superposition entre agents
  - Champ de vision circulaire avec rayon paramétrable.
  - Angle mort du champ de vision avec rayon paramétrable.
- Simulation sur Unity.
  - Taille de l’environnement : 7.0 x 7.0 unité
  - Taille de l’agent : 0.08 unité.
  - Nombre d’agents : 40
  - Le visuel des agent est un cylindre de couleur noire.
  - L’environnement est de couleur blanche.
  - Le point de vue est : vue de dessus.
  - Les clips sont joués à 60 fps.
  - Il y a une pause entre chaque clip, le participant doit lancer le clip suivant en appuyant sur la touche "espace".

#### 4.4 Procédure

L’expérimentation se tiendra en 3 phases : une première de familiarisation, une seconde de manipulation et une troisième phase de débriefing.

##### **Familiarisation :**

- Accueil du participant.
- Présentation au participant de l’expérimentation, signature du formulaire de consentement.
- (Facultatif) Questionnaire pour recueillir des informations sur le participant.
- Explication du concept de fracture au participant, accompagné de deux clips en exemple.
- Explication de la tâche à réaliser au participant : Choisir la bonne réponse à chaque clip, et valider cette réponse le plus rapidement possible dès que le participant est sûr de lui.
- Présentation des 3 touches du clavier utilisées : espace, 'a', 'p'.

##### **Manipulation :**

- Début de l’expérimentation, les clips vont défiler les uns après les autres et le participant doit répondre à chaque clip, le plus rapidement possible dès qu’il est sûr de lui

##### **Débriefing :**

- Phase de discussion sur quelques clips choisis. Enregistrement audio (et capture d’écran) de la discussion (si autorisé par le participant).
- Remplissage du questionnaire de fin.
- Fin de l’expérimentation, temps de réponses aux éventuelles questions du participant.

## 4.5 Caractéristiques des clips

Les clips présentés aux participants sont caractérisés comme tel :

- Les clips sont d'une durée variable (6s - 50s)
- Les clips seront classés en 4 catégories en fonction du comportement de l'essaim : densification, expansion, flocking, swarming.
- Il y a au total 48 clips :
  - 12 clips de chaque catégorie
  - 6 clips avec fracture et 6 clips sans fracture dans chaque catégorie (ce qui fait au total 24 avec fractures et 24 sans fractures)
- La durée des 48 clips est d'environ 19 minutes.
- Les clips sont obligatoirement visionnés jusqu'à la fin par tous les participants.
- Les clips sont ordonnés aléatoirement pour chaque participant.
- Il y a 8 clips supplémentaires pour la phase de discussion (3 sans fractures, 5 avec).

## 4.6 Variables

### Observables de l'expérimentateur :

- Choix du participant
- Horodatage de la réponse du participant
- La réflexion à voix haute du participant
- Le questionnaire de fin

### Clips :

- Les catégories des clips :
  - densification : Les agents de l'essaim se rapprochent les uns des autres, provoquant ainsi une augmentation de la densité de l'essaim.
  - expansion : Opposé de la densification. Les agents de l'essaim s'éloignent les uns des autres, provoquant ainsi une diminution de la densité de l'essaim.
  - flocking : Les agents se déplacent à la même vitesse et avec la même orientation. On peut observer les agents se déplaçant en groupe, de façon coordonnée.
  - swarming : Opposé du flocking. Les agents se déplacent sans aucune coordination, en allant dans toutes les directions et avec des vitesses différentes. Les agents restent cependant en groupe.
- Métrique "DistanceVisibilityScore" : Cette métrique calcule la distance entre deux groupes fracturés par rapport à la distance inter-agent de l'essaim principal.
- Métrique "SeparationSpeed" : Cette métrique calcule la vitesse d'éloignement du groupe fracturé par rapport à l'essaim principal au moment de la fracture.

## 4.7 Hypothèses opérationnelles

1. (*théorique*) Les humains ne perçoivent pas tout le temps les fractures des essais.
2. (*théorique*) Les humains ne perçoivent pas les fractures avant que celles-ci n'apparaissent.
3. (*théorique*) Les humains ne perçoivent pas le moment/seuil exact de la fracture de l'essai.
4. (*théorique*) Les humains perçoivent mieux les fractures quand la distance inter groupes est supérieur à la distance intra groupe.
  - (*opérationnelle*) Les taux de bonnes réponses par clip des participants exprimeront une meilleure perception des fractures quand la meilleure valeur de "DistanceVisibilityScore" sur les clips est importante.
5. (*théorique*) Les humains prévoient mieux les fractures quand la distance inter groupes est supérieur à la distance intra groupe au moment de la fracture.
  - (*opérationnelle*) Les taux de bonnes réponses par clip et le moment de la réponse des participants exprimeront une meilleure prévision des fractures quand la valeur de "DistanceVisibilityScore" au moment de la fracture sur les clips est importante.
6. (*théorique*) Les humains perçoivent mieux les fractures quand le groupe responsable de la fracture s'éloigne rapidement du groupe principal.
  - (*opérationnelle*) Les taux de bonnes réponses par clip des participants exprimeront une meilleure perception des fractures quand la valeur de "SeparationSpeed" sur les clips est importante.
7. (*théorique*) Les humains prévoient mieux les fractures quand le groupe responsable de la fracture s'éloigne rapidement du groupe principal.
  - (*opérationnelle*) Les taux de bonnes réponses par clip et le moment de la réponse des participants exprimeront une meilleure prévision des fractures quand la valeur de "SeparationSpeed" au moment de la fracture sur les clips est importante.
8. (*théorique*) Certaines catégories provoquent chez un humain un jugement de prédiction différent des autres catégories (peut importe la réponse).
  - (*opérationnelle*) Le taux de réponses données avant la fracture sera plus importante dans une catégorie que les autres.
  - (*opérationnelle*) Le taux de mauvaises réponses données avant la fracture sera plus importante dans une catégorie que les autres.
9. (*théorique*) Certaines catégories sont plus difficiles à percevoir pour un humain.
  - (*opérationnelle*) Le taux de bonne réponses de la catégorie "expansion" exprimera un moins bon résultat qu'avec les autres catégories.
10. (*théorique*) Certaines catégories sont plus facile à prédire pour un humain.
  - (*opérationnelle*) Le taux de bonne réponses avant la fracture de la catégorie "densification" exprimera un meilleur résultat qu'avec les autres catégories.

## 4.8 Recueil des données

- Horodatage de la réponse du participant (automatiquement au moment de la réponse du participant)
- Le résultat de la réponse du participant (si il ne répond pas la réponse est considérée comme étant: non il n'y a pas de fracture) (automatiquement au moment de la réponse du participant)

- Audio + capture d'écran : Thinking Out Loud, lors d'une phase de discussion en visionnant des clips pré-sélectionnés (afin de ne pas impacter le résultat des participants [Hertzum et al., 2009] car on leur demande en priorité d'être performant).
- Un questionnaire subjectif

## 4.9 Analyse des résultats

- Calcul du taux de bonnes réponses total (avec fracture, sans fracture et les deux)
- Calcul du taux de bonnes réponses par participant (avec fracture, sans fracture et les deux)
- Calcul du taux de bonnes réponses par catégorie de clips (avec fracture, sans fracture et les deux)
- Calcul du taux de réponses avant la fracture total (avec fracture, sans fracture et les deux)
- Calcul du taux de réponses avant la fracture par participant (avec fracture, sans fracture et les deux)
- Calcul du taux de réponses avant la fracture par catégorie (avec fracture, sans fracture et les deux)
- Calcul du taux de bonnes réponses avant la fracture total (avec fracture, sans fracture et les deux)
- Calcul du taux de bonnes réponses avant la fracture par participant (avec fracture, sans fracture et les deux)
- Calcul du taux de bonnes réponses avant la fracture par catégorie (avec fracture, sans fracture et les deux)
- Traitement statistique : Anova à 2 voies (fractures x catégorie)
- Analyse de l'audio des participants (récupération des termes les plus utilisés, potentiellement caractérisant des éléments essentiels de la perception des essaims)
- Analyse des questionnaires

## References

- [Hertzum et al., 2009] Hertzum, M., Hansen, K. D., and Andersen, H. H. (2009). Scrutinising usability evaluation: does thinking aloud affect behaviour and mental workload? *Behaviour & Information Technology*, 28(2):165–181.