

Attention, Perception, & Psychophysics

Detecting single-target changes in multiple object tracking: The case of peripheral vision

Journal:	<i>Attention, Perception, & Psychophysics</i>
Manuscript ID	PP-ORIG-15-140.R2
Manuscript Type:	Original Manuscript
Date Submitted by the Author:	n/a
Complete List of Authors:	Vater, Christian; University of Bern, Kredel, Ralf; University of Bern, Hossner, Ernst-Joachim; University of Bern,
Keywords:	Attention: object-based, eye movements and visual attention, visual search

1
2
3 1 Running Head: DETECTING TARGET CHANGES WITH PERIPHERAL VISION
4
5
6
7
8
9
10
11
12
13

14 6 **Detecting single-target changes in multiple object tracking:**

15
16 7 **The case of peripheral vision**
17
18
19 8

20 9 Christian Vater¹, Ralf Kredel¹, & Ernst-Joachim Hossner¹

21
22
23 10 ¹University of Bern, Switzerland
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

13 **Corresponding author:**

14 Christian Vater

15 University of Bern

16 Bremgartenstrasse 145

17 CH-3012 Bern

18 Switzerland

19 Tel: +41 31 631 52 07

20 Mail: Christian.Vater@ispw.unibe.ch

22 Names and mail addresses of co-authors:

23 Ralf Kredel, Ralf.Kredel@ispw.unibe.ch

24 Ernst-Joachim Hossner, Ernst.Hossner@ispw.unibe.ch

Abstract

In the current study it is investigated whether peripheral vision can be used to monitor multiple moving objects and to detect single-target changes. For this purpose, in Experiment 1, a modified MOT setup with a large projection and a constant-position centroid phase had to be checked first. Classical findings regarding the use of a virtual centroid to track multiple objects and the dependency of tracking accuracy on target speed could be successfully replicated. Thereafter, the main experimental variations regarding the manipulation of to-be-detected target changes could be introduced in Experiment 2. In addition to a button press used for the detection task, gaze behavior was assessed using an integrated eye-tracking system. The analysis of saccadic reaction times in relation to the motor response shows that peripheral vision is naturally used to detect motion and form changes in MOT because the saccade to the target occurred after target-change offset. Furthermore, for changes of comparable task difficulties, motion changes are detected better by peripheral vision than form changes. Findings indicate that capabilities of the visual system (e.g., visual acuity) affect change detection rates and that covert-attention processes may be affected by vision-related aspects like spatial uncertainty. Moreover, it is argued that a centroid-MOT strategy might reduce the amount of saccade-related costs and that eye-tracking seems to be generally valuable to test predictions derived from theories on MOT. Finally, implications for testing covert attention in applied settings are proposed.

Keywords

covert attention; perception; motor control; saccadic latency; eye-tracking, sports

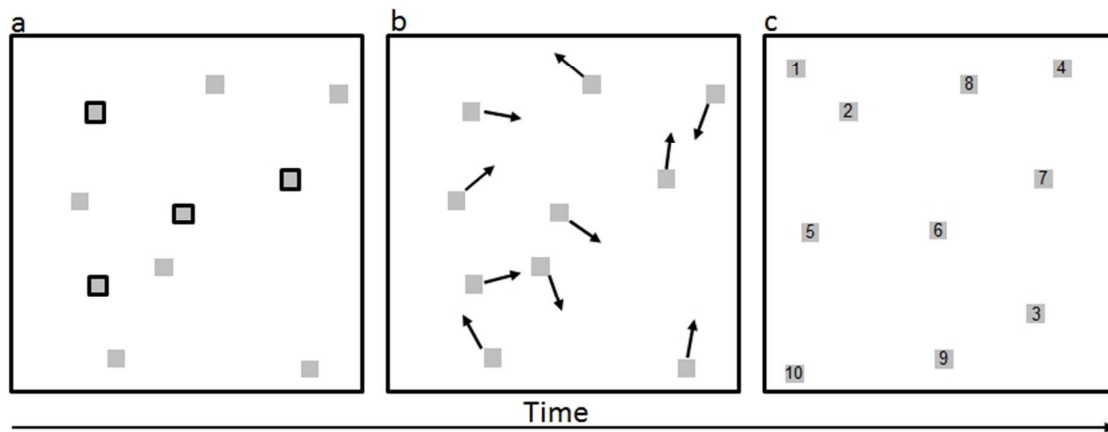
47 Introduction

48 Peripheral vision allows us to detect natural objects at large eccentricities at up to
49 70.5° eccentricity level (Thorpe, Gegenfurtner, Fabre-Thorpe, & Bülthoff, 2001) and to cate-
50 gorize natural scenes at 70° eccentricity with high accuracy (Boucart, Moroni, Thibaut, Szaf-
51 farczyk, & Greene, 2013) even under crowded conditions (Li, VanRullen, Koch, & Perona,
52 2002). Neuroanatomically, the capability to process visual information in the far periphery
53 seems to be controlled by distinct, specialized cortical networks (for a recent review see Yu,
54 Chaplin, & Rosa, 2015). Furthermore, in comparison to foveal vision, peripheral vision fea-
55 tures higher temporal resolution (Hartmann, Lachenmayr, & Brettel, 1979) and superior mo-
56 tion detection (Finlay, 1982; McKee & Nakayama, 1984). Thus, while foveal vision with its
57 high spatial resolution is important for extracting detailed information, a crucial role for pe-
58 ripheral vision seems to refer to the processing of changing aspects in the visual environment
59 (for a review, see Strasburger, Rentschler, & Jüttner, 2011).

60 Hence, a monitoring task on the detection of motion and form changes in the periphery
61 can be regarded as a paradigmatic example of making use of peripheral vision. In an applied
62 sense, those tasks can be found in team sports, where it is often necessary to perceive a num-
63 ber of moving objects across a large functional field of view (Davids, 1984). This applies, for
64 example, in offside decision-making in soccer, where the assistant referee has to monitor the
65 player in ball possession as well as the players at the offside line. In this situation, experts
66 show a gaze strategy in which they anchor their gaze on the offside line while perceiving the
67 player in ball possession peripherally (Catteeuw, Helsen, Gilis, Van Roie, & Wagemans,
68 2009). Furthermore, Williams and Davids (1998) were able to empirically show that, in a
69 three-versus-three anticipation task in soccer, experts extract a great amount of information
70 regarding players' positions and movements with peripheral vision. The importance of track-
71 ing multiple objects in team sports was also underlined by Faubert and Sidebottom (2012),

72 who described the necessity for soccer goalkeepers to simultaneously track the ball, team-
73 mates, and opponents while keeping a large and dynamic visual field under surveillance.

74 In experimental psychology, the demands imposed from ecologically valid situations
75 are generally studied with *multiple-object tracking* (MOT), a task, which was introduced by
76 Pylyshyn and Storm (1988) to study visual attention. The task requires tracking a certain
77 number of targets amidst identically looking distractors for a pre-defined amount of time. Be-
78 fore the objects begin to move on the computer screen, the targets are briefly highlighted, then
79 all objects start to move in a (quasi-)random fashion, and finally, after targets have stopped,
80 participants are supposed to recall the targets (see Figure 1 in which stimulus material of the
81 current study is illustrated).



82
83 *Figure 1.* Experimental paradigm used in Experiment 1 and 2. After designating the four targets with red frames
84 (a), the frames disappear and all stimuli move in quasi-random fashion for 6 s (b). At the end of the motion
85 phase, participants have to recall the targets by naming the respective numbers that are projected onto the objects
86 (c).

87 Theoretically, several approaches have been brought forward to explain tracking per-
88 formance in MOT, namely, (a) the FINST (Fingers of INSTantiation) model by Pylyshyn and
89 Storm (1988), (b) the grouping model by Yantis (1992), and (c) the multifocal attention model
90 by Cavanagh and Alvarez (2005). After the brief descriptions of these approaches, we will

1
2
3 91 come back to the question how specifics of peripheral vision can be related to these concepts
4
5 92 in particular and to MOT performance in general.
6
7

8 93 (a) In their MOT studies, Pylyshyn and Storm (1988) were mainly interested on how
9
10 94 visual attention can be shifted independent of eye movements and on whether locations are
11
12 95 scanned serially or in parallel. According to their feature-binding-based FINST model, partic-
13
14 96 ipants track multiple objects in parallel while each object possesses an individual internal ref-
15
16 97 erence at a pre-attentive stage. The authors describe a serial tracking algorithm in which target
17
18 98 locations (instead of motion vectors) are stored. In a related experiment, motor response la-
19
20 99 tencies to flashes on targets were measured as a function of the number of targets. Latencies
21
22 100 significantly increased from one target to five targets which could be regarded as an indicator
23
24 101 that participants serially track the targets (Pylyshyn & Storm, 1988). However, in a second
25
26 102 experiment, to test the serial-scanning algorithm, participants had to indicate whether a flash
27
28 103 occurred on a target, distractor, or another location. As the observed performance turned out
29
30 104 to be much higher than the algorithmically predicted one, Pylyshyn and Storm finally con-
31
32 105 cluded that targets are rather tracked in parallel than serially.
33
34
35
36

37 106 (b) Yantis (1992) alternatively proposed that participants keep track of the targets by
38
39 107 making use of a virtual polygon formed out of the targets. Two stages are predicted by this
40
41 108 model. The first one, the group formation stage, is influenced by Gestalt laws of grouping and
42
43 109 can be characterized as pre-attentive, automatic, and stimulus-driven and, thereby, compara-
44
45 110 ble to the model introduced by Pylyshyn and Storm (1988). However, in contrast to the
46
47 111 FINST model, the second stage that concerns group maintenance is characterized as goal-
48
49 112 directed, effortful, and attention-demanding. Furthermore, a continuous updating of target
50
51 113 representations is assumed and discussed in the context of mental rotations. Yantis (1992)
52
53 114 experimentally tested these predictions by applying different Gestalt laws to the MOT task.
54
55 115 For example, when five targets either rotated in the same direction (rigid condition) or three
56
57
58
59
60

1
2
3 116 targets in one and two targets in the other direction (nonrigid condition), response accuracies
4
5 117 were higher in the rigid condition, presumably because targets followed the law of common
6
7 118 fate. Furthermore, Zhong Ma, Wilson, Liu, and Flombaum (2014) showed that, rather than
8
9 119 motion extrapolation, the recently observed object position seems to be used for tracking.
10
11 120 This finding can be interpreted as support for the effortful updating process in the Yantis
12
13 121 model because, on the basis of the continues updating, no motion extrapolation would be ex-
14
15 122 pected (for an overview on mixed results on motion extrapolation, however, see Zhong et al.,
16
17 123 2014).

20
21 124 (c) Finally, Cavanagh and Alvarez (2005) proposed a multifocal attentional mecha-
22
23 125 nism for MOT that is based on the assumption that attention is split to the targets while each
24
25 126 selection channel comprises a position tracker and a stream of object features. This model is
26
27 127 underpinned, among others, by the results presented by Meyerhoff, Papenmeier, Jahn, and
28
29 128 Huff (2013) who were able to show that unexpected changes in targets do affect tracking per-
30
31 129 formance while changes in distractors have no effect even if these changes were in the line of
32
33 130 sight. To examine whether motion information are actually used for tracking, Huff and col-
34
35 131 leagues introduced motion-texture information on the targets moving either in the same or in a
36
37 132 different direction as a target, thereby, not confounding spatiotemporal information of the
38
39 133 objects with motion information (Huff & Papenmeier, 2013; Meyerhoff, Papenmeier, & Huff,
40
41 134 2013; St.Clair, Huff, & Seiffert, 2010). Among others, these studies show that tracking per-
42
43 135 formance declines when the difference between object direction and texture direction increas-
44
45 136 es (St.Clair et al., 2010) and that the availability and reliability of spatiotemporal and feature
46
47 137 information is considered by participants (Meyerhoff, Jahn & Huff, 2014). As motion and
48
49 138 feature information seem to be relevant for tracking performance, these results are perfectly in
50
51 139 line with the multifocal attention theory of MOT.
52
53
54
55
56
57
58
59
60

1
2
3 140 In the three models sketched so far, the theoretical focus was laid on questions on vis-
4
5 141 ual attention, mostly without applying eye-tracking technology to the respective studies. In
6
7 142 this vein, Fehd and Seiffert (2008) demonstrated that multiple target objects are tracked
8
9 143 amongst distractors with a center-looking strategy relying on a virtual centroid, which is de-
10
11 144 fined as the center of mass of the polygon formed by the targets. Furthermore, they showed
12
13 145 that this strategy is independent of speed and object size and that tracking performance is
14
15 146 highest when a center-target switching strategy is used in which gaze is shifted back and forth
16
17 147 between the centroid and the targets. Most interestingly in regards to the observed gaze be-
18
19 148 havior, slower stimuli did not lead to more saccades so that a saccade-avoiding strategy was
20
21 149 not evident. On the one hand, this could have been expected since increased effects of crowd-
22
23 150 ing are usually observed at higher stimulus speed (Franconeri, Lin, Pylyshyn, Fisher, & Enns,
24
25 151 2008), which would lead to the necessity to saccade more frequently to targets because of
26
27 152 potential collisions (Elfanagely Haladjian, Aks, Kourtev, & Pylyshyn, 2011; Landry, Sheri-
28
29 153 dan, & Yufik, 2001). On the other hand, while saccading helps updating the position of single
30
31 154 targets (Landry et al., 2001), the costs of eye movements increase with higher stimulus speed
32
33 155 because, when executing saccades, objects move far if stimulus speed is high (Huff, Pa-
34
35 156 penmeier, Jahn, & Hesse, 2010). Additionally, saccadic suppression of information pro-
36
37 157 cessing (about 75 ms before saccade onset to 50 ms after saccade offset) makes the updating
38
39 158 process even more difficult (Diamond, Ross, & Morrone, 2000). In this context, the role of
40
41 159 “rescue saccades” in MOT has been examined by Zelinsky and Todor (2010) who found that
42
43 160 anticipatory saccades are driven by the potential for losing track of a target when it is close to
44
45 161 a barrier occluder or a distractor. It could, therefore, be that participants preferably use their
46
47 162 peripheral vision to monitor target positions while “anchoring” their gaze on the virtual cen-
48
49 163 troid (Fehd & Seiffert, 2010; Zelinsky & Neider, 2008) and initiate saccades just before the
50
51 164 distance of a target to other objects becomes too small (Zelinsky & Todor, 2010).
52
53
54
55
56
57
58
59
60

1
2
3 165 When relating the attention-related explanations of MOT behavior (see above: a-c) to
4
5 166 the just sketched gaze studies, it seems likely that covert attention is used for tracking as gaze
6
7 167 is apparently on a virtual centroid in large part. Thus, to examine the actual location of atten-
8
9 168 tion, target changes has been included in experimental studies on MOT. In this branch of re-
10
11 169 search, target movements (motion direction) or target properties (color, form) were manipu-
12
13 170 lated to investigate, on the one hand, whether target-motion information are used for tracking
14
15 171 and, on the other hand, how attention is distributed to targets and distractors. Regarding the
16
17 172 use of motion information, Meyerhoff et al. (2013) studied the effect of random motion-
18
19 173 direction changes (deviation of up to 60 ° to the left or right) of either targets, distractors, or
20
21 174 both. The result, that an *unexpected* change of a target but not of a distractor vector impaired
22
23 175 performance, shows that, in line with previous findings (e.g., Fencsik, Klieger, & Horowitz,
24
25 176 2007), target-motion information are crucial for MOT. However, if changes are *expected* and
26
27 177 participants have to respond to them, it is possible to figure out where attention is located dur-
28
29 178 ing MOT. Therefore, Bahrami (2003) introduced color (red, green, blue) and form (“T”, “L”,
30
31 179 “+”) changes in targets and distractors. He found that color and form changes of the target
32
33 180 were more often detected than changes of distractors, again showing that attention is attracted
34
35 181 by targets. Moreover, color changes were identified more frequently (approx. 80 %) than form
36
37 182 changes (approx. 55 %). In a closely related study, Sears and Pylyshyn (2000) induced either
38
39 183 a target or a distractor form change in 50 % of the trials (shape of a “seven-segment box fig-
40
41 184 ure eight” becoming either an “E” or an “H”). They showed that target changes were detected
42
43 185 much faster than distractor changes, while a higher number of distractors increased response
44
45 186 times when target or distractor changes were evident, underlining the finding that attention is
46
47 187 rather attracted by targets than by distractors. This result is also supported by electrophysio-
48
49 188 logical measures, for example, by Drew, McCollough, Horowitz and Vogel (2009) who
50
51 189 measured event-related potentials (ERPs) while participants had to track two targets amidst
52
53
54
55
56
57
58
59
60

1
2
3 190 four moving distractors and four stationary distractors. Task-irrelevant white square probes
4
5 191 were presented randomly at the target, distractor, empty space or a stationary target. As ex-
6
7 192 pected, target probes showed the greatest ERP-response while weaker responses were ob-
8
9 193 served for distractors and weakest for background and stationary targets. Hence, behavioral
10
11 194 and electrophysiological results confirmed that spatial attention is more on targets than on
12
13 195 other locations.

14
15
16
17 196 Summing up, regarding theoretical approaches to explain MOT performance, recent
18
19 197 research rather supports the multifocal attention theory (c) as target motion and feature infor-
20
21 198 mation are apparently used in the tracking process. Furthermore, the centroid strategy re-
22
23 199 vealed by eye-tracking studies suggests that MOT performance is mainly based on covert at-
24
25 200 tention which, in turn, is particularly allocated to the targets.

26
27
28 201 As it becomes obvious from this summary, despite the fact that eye-tracking technolo-
29
30 202 gy has been applied, results revealed in MOT studies were mostly discussed from an attention
31
32 203 perspective and rarely from a vision perspective. This comes as a surprise because many find-
33
34 204 ings could also be explained by the basal properties of the peripheral visual system, in particu-
35
36 205 lar, by the decreasing visual acuity but increasing contrast sensitivity of the peripheral retina.
37
38 206 For example, when Bahrami (2003) reports that target color changes are detected more often
39
40 207 than form changes this could also be traced back to luminance and contrast sensitivity differ-
41
42 208 ences of the retina such that foveal vision might be advantageous in detecting form changes
43
44 209 due to its higher spatial resolution (Gralla, 2007). Beside this, the higher temporal resolution
45
46 210 (Hartmann et al., 1979) and the superior motion-detection capability of the peripheral visual
47
48 211 system (Finlay, 1982; McKee & Nakayama, 1984) would also predict high detection rates for
49
50 212 motion changes in MOT. In this regard, it is really surprising that, up to date, no attempt has
51
52 213 been made to implement a motion-change detection task to examine the role of peripheral
53
54 214 vision in MOT. In a respective study, it would seem crucial to control gaze behavior as in
55
56
57
58
59
60

1
2
3 215 previous experiments (e.g., Sears & Pylyshyn, 2000) participants were only instructed to track
4
5 216 target objects while keeping their eyes on the fixation cross without including eye-tracking
6
7 217 technology. Consequently, up to now, it is completely unknown which natural gaze strategy
8
9 218 is executed in MOT to detect target changes in a setting in which no explicit instructions on
10
11 219 vision are given. Hence, exactly this empirical question will be addressed in the following
12
13 220 experiments.

14
15
16
17 221 From an applied perspective, it should be added that these experiments promise the
18
19 222 derivation of helpful advices for the training of gaze strategies, for example, in team-sport
20
21 223 situations as they were sketched in the beginning of this paper. In this regard, form-changes in
22
23 224 MOT can be associated with posture-related changes (e.g., a basketball player indicating with
24
25 225 his hands where he wants to receive the ball) and motion-changes in MOT with movement-
26
27 226 related changes (e.g., detecting a penetration of an opposing player to the basket). As the reli-
28
29 227 ability of such recommendations grows with the external validity of the experimental setting,
30
31 228 it seems worthwhile to study those changes by displaying the MOT task on a large screen to
32
33 229 overcome the limitation of a restricted field of vision when examining the relevance of pe-
34
35 230 ripheral vision. Hence, in a series of two studies, Experiment 1 aims to answer the question
36
37 231 whether classical MOT findings can be replicated under those novel experimental conditions.
38
39 232 Based on the results of Experiment 1, the main research question will be approached in Ex-
40
41 233 periment 2 in which the focus will be on the case of peripheral vision in monitoring multiple
42
43 234 moving objects while detecting changes in the periphery.

44
45
46
47
48
49 235

50 51 236 **Experiment 1**

52
53
54 237 Experiment 1 was designed as a replication study, in particular regarding the MOT results
55
56 238 reported by Fehd and Seiffert (2008, 2010) on the effects of object speed on response accura-

1
2
3 239 cy, on the one hand, and on gaze behavior, on the other hand. To increase the external validity
4
5 240 of the study, a large screen was used to display the MOT task. In addition, a novel manipula-
6
7 241 tion detail was introduced in anticipation of the main research question on form versus motion
8
9 242 changes in Experiment 2. In this experiment, it seems to be crucial to manipulate the motion
10
11 243 of targets and distractors in such a way that form- or motion-related changes occur inde-
12
13 244 pendently of (uncontrolled) movements of the centroid as well as of (uncontrolled) eccentrici-
14
15 245 ties of the to-be-detected event. For this reason, a stationary phase of the centroid was applied
16
17 246 in which the respective event occurred at a defined eccentricity (under the assumption that the
18
19 247 centroid was actually focused). In spite of the larger display and the introduction of a station-
20
21 248 ary-centroid phase, a successful replication of the findings reported by Fehd and Seiffert
22
23 249 (2008, 2010) was expected. Hence, it was predicted, that response accuracy decreases as a
24
25 250 function of object speed and that participants prefer a centroid-tracking strategy independent
26
27
28
29
30 251 of speed conditions.

31 32 252 **Methods**

33
34
35 253 *Participants.* 14 sport science students (7 females and 7 males; age: 21.7 ± 1.3 years)
36
37 254 participated in the experiment and received course credits in return. They had self-reported
38
39 255 normal or corrected-to-normal vision and were unaware of the research question. The experi-
40
41 256 ment was undertaken in accordance with the Declaration of Helsinki.

42
43
44
45 257 *Stimuli.* Motion paths of the objects (10 white squares, 35 mm x 35 mm corresponding
46
47 258 to $1^\circ \times 1^\circ$ of visual angle) were calculated with Matlab and then imported to Autodesk 3ds
48
49 259 Max (2014) to render single video trials. Stimuli were presented within a quadratic frame
50
51 260 (white line of 25 mm width, 1.40 m x 1.40 m corresponding to $40^\circ \times 40^\circ$ of visual angle) on a
52
53 261 black background. All squares appeared in (quasi-)randomized starting positions precluding
54
55
56 262 overlaps with each other. Each trial began with a trial number, followed by a stationary
57
58
59
60

1
2
3 263 presentation of ten squares, including four targets that were highlighted by red frames (line of
4
5 264 15 mm width; frame and stimulus together covering an area of $1.7^\circ \times 1.7^\circ$ of visual angle).
6
7 265 After 2 s, target cues disappeared, and all stimuli accelerated over one second to the final
8
9 266 speed of either $6^\circ/\text{s}$ or $9^\circ/\text{s}$ or $12^\circ/\text{s}$, which was retained for 4 s. After a subsequent decelera-
10
11 267 tion phase of one further second, all squares stopped, so the overall motion time was 6 s per
12
13 268 trial.

14
15
16
17 269 During the motion time, a repulsion mechanism was used to redirect a square whenever
18
19 270 er the distance to the bordering frame or to the next square fell below a certain threshold
20
21 271 (35 mm corresponding to 1° of visual angle). To ensure independence of the findings of cen-
22
23 272 troid-related motions (as well as of different eccentricities of target changes in Experiment 2),
24
25 273 targets' paths were manipulated such that the centroid was forced to stay at a constant posi-
26
27 274 tion for 0.5 s. For this purpose, the path of one distractor was calculated in such a way that it
28
29 275 collided with one target at one of the three possible points in time (3.0 s, 3.5 s, 4.0 s after mo-
30
31 276 tion onset) and that all targets moved after the collision in a concerted way in angles of 90°
32
33 277 difference to each other (e.g., 10° , 100° , 190° , and 280° in relation to the frame's baseline).
34
35 278 Due to the same speed of the targets, this specification resulted in a stationary position of the
36
37 279 centroid of the polygon formed by the four targets. Over the 0.5 s following the critical colli-
38
39 280 sion, the five remaining distractors were allowed to move randomly; however, it was ensured
40
41 281 that no collision of a target with another object or the bordering frame occurred within this
42
43 282 time window of 0.5 s. Before and after this interval, all objects were allowed to collide with
44
45 283 one another or with the bordering frame without any restrictions.

46
47
48
49
50
51 284 On the basis of these constraints, 14 mother trials were created, which differed on the
52
53 285 position of the crucial collision, the resulting position of the stationary centroid, and the con-
54
55 286 certed motion direction of the target during the stationary-centroid phase. From the mother
56
57 287 trials, a total of 126 trials ($= 14 \times 3 \times 3$) was derived by starting the stationary-centroid phase
58
59
60

1
2
3 288 at three different onsets (either 3.0 s, 3.5 s, or 4.0 s after motion onset) and by accelerating
4
5 289 targets and distractors to three final speeds (either 6 °/s, 9 °/s, or 12 °/s). To balance trials of
6
7 290 different speeds within blocks, each block contained five trials of each speed, which were
8
9 291 presented in a randomized order (rendering by MAGIX Video Pro X3). Since a single session
10
11 292 of 60 min allowed for the presentation of a maximum of 9 blocks and participants' capability
12
13 293 to keep focused on the quite demanding task seemed to be limited, exactly this number of 15-
14
15
16 294 trial blocks was employed. Consequently, in the resulting total of 135 trials, 9 of the 126 trials
17
18 295 derived from the mother trials were presented twice (i.e., three per stationary-centroid phase
19
20
21 296 onset and speed).

22
23 297 *Apparatus.* A binocular eye-tracking system (EyeSeeCam, 220 Hz) was used to assess
24
25 298 the vertical and horizontal rotations of both eyes via infrared reflections from the pupil and
26
27 299 the cornea (measurement accuracy: 0.5 ° of the visual angle with a resolution of 0.01 ° RMS
28
29 300 within 25 ° of the participant's field of view). The EyeSeeCam (ESC) is connected to a Mac-
30
31 301 Book Pro via a 20 m fiber-optic Fire Wire link (GOF-Repeater 800, Unibrain), which is
32
33 302 stored in a bum bag. Since the eye-tracker is synchronized with a 10-camera VICON-T20
34
35 303 system that tracks retro-reflective markers attached to the ESC, a three-dimensional gaze vec-
36
37 304 tor in a laboratory frame of reference can be calculated (Kredel et al., 2011). This gaze vector
38
39 305 is updated every 5 ms and allows for relating the participant's current gaze to either the dis-
40
41 306 played targets or distractors or the centroid derived from the current positions of the targets.
42
43
44
45

46 307 The ESC was calibrated at the beginning of each test session. For this purpose, partici-
47
48 308 pants had to consecutively fixate five dots of a regular grid with a distance of 8.5 ° of visual
49
50 309 angle between the dots (Kredel et al., 2011). The ESC was recalibrated before each test block
51
52 310 whenever the point of gaze deviated more than 0.5 ° of visual angle from one of the dots in
53
54 311 the calibration grid. Stimuli were back-projected (projector: InFocus IN5110) onto a large
55
56 312 screen (height: 1.87 m; width: 3.01 m), while the quadratic frame for the MOT task covered
57
58
59
60

1
2
3 313 an area of 1.40 m x 1.40 m in the middle of the screen. Microsoft Windows Media Player was
4
5 314 used to playback the video trials. The gathered data were analyzed using Mathworks
6
7 315 MATLAB 2013a. Further statistical analyses were conducted with IBM SPSS Statistics 22.
8
9

10 316 *Procedure.* Participants were tested individually in a single one-hour session in the in-
11
12 317 stitute's sensorimotor laboratory. After reading the general information about the study and
13
14 318 agreeing to participate, the eye-tracking system was fitted. Subsequently, participants were
15
16 319 positioned at 2 m distance from the screen for reading the displayed instructions about the to-
17
18 320 be-solved task. The task was to recall the targets that were cued at the beginning by naming
19
20 321 the respective numbers that were projected onto the targets at the end of each trial. After-
21
22 322 wards, the ESC calibration routine was conducted. After each trial, participants had to recall
23
24 323 the targets by naming the respective numbers that were displayed on the ten squares. Partici-
25
26 324 pants' verbal decisions were recorded in writing by an experimenter. No augmented feedback
27
28 325 on the responses' correctness was given after the trials.
29
30
31
32

33 326 *Gaze analysis.* For the assessment of the location of gaze, a region-of-interest method
34
35 327 was applied. For this purpose, eleven regions of interest were defined as circles with a diame-
36
37 328 ter of 5 ° over each square as well as over the centroid, which was calculated as the center of
38
39 329 mass of the polygon formed by the four targets (Fehd & Seiffert, 2010). Whenever the current
40
41 330 gaze vector hit a region of interest, the respective frame was counted for the related object.
42
43 331 This means that, in cases of overlapping regions, an allocation to two or even more objects
44
45 332 was possible. The number of counted frames per region was then divided by the total number
46
47 333 of frames in the trial (i.e., by 1200 frames for the 6 s of motion time) resulting in percentages
48
49 334 of gaze-allocation time for each region of interest, a variable that has been termed "gaze over-
50
51 335 lap time" by Fehd and Seiffert (2010).
52
53
54
55
56
57
58
59
60

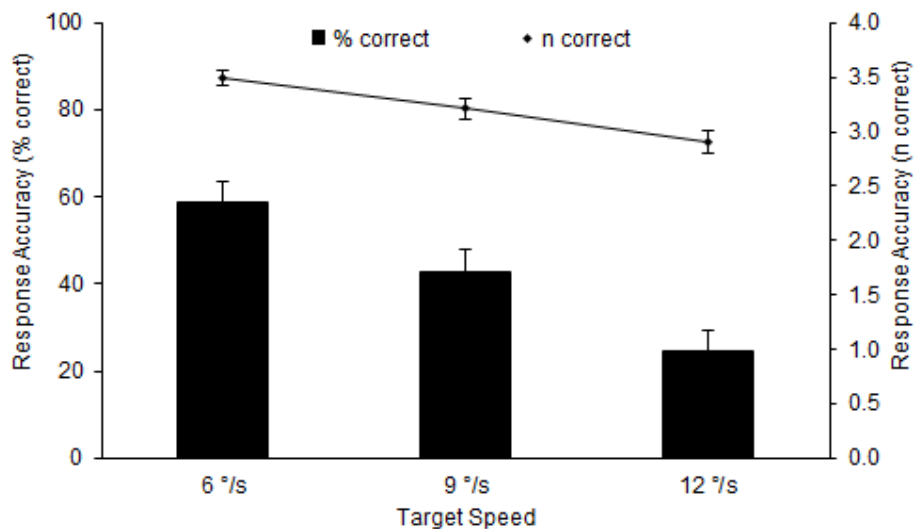
1
2
3 336 *Measures.* The first group of dependent variables refers to the correctness of the 45 re-
4
5 337 sponses per speed condition, which can be expressed either as the average number of correctly
6
7 338 recalled targets (n correct) or as the percentage of trials in which all four targets could be suc-
8
9 339 cessfully recalled (% correct). Only successful trials in the latter sense were included in the
10
11 340 aggregation of gaze-related variables, as in cases when just a subset of targets was correctly
12
13 341 identified, participants have probably lost one of the targets somewhere over the trial such
14
15 342 that at the point of time the calculation of the virtual centroid would have lost its validity.
16
17 343 Thus, the above-defined variable of gaze-overlap time was calculated for trials with perfect
18
19 344 target recall only (% of total time). Since no specifics of a certain target or a certain distractor
20
21 345 were to be expected, gaze-overlap times were calculated as the averages of one centroid-
22
23 346 related, four target-related, and six distractor-related values, respectively.

24
25
26
27
28 347 All dependent variables were analyzed with repeated-measures ANOVAs with the
29
30 348 three speed conditions as the within-subject factor, in the case of the gaze-allocation analysis
31
32 349 completed by a second within-subject factor for the three regions of interest (centroid, target
33
34 350 or distractor). As the only reason for including three motion-onset conditions was to prevent
35
36 351 participants from learning of the otherwise to-be-anticipated onset-time of the respective
37
38 352 event and as no theory-related predictions can be derived for this variable, the stationary-
39
40 353 centroid phase onset was not included as a factor in the ANOVA calculations. Significant
41
42 354 interaction effects were further analyzed with planned t-tests. Based on the results reported by
43
44 355 Fehd and Seiffert (2010), the alpha level for tests on differences was a priori set to $\alpha = .05$. A
45
46 356 posteriori effect sizes were computed as partial eta squares (η_p^2), and in cases of non-
47
48 357 significant tests, the power ($1-\beta$) was calculated. According to Fehd and Seiffert (2010), high-
49
50 358 est response accuracies were expected for the 6 °/s speed condition, followed by the 9 °/s and
51
52 359 the 12 °/s conditions. Furthermore, gaze overlap was expected to be highest for the centroid,
53
54
55 360 followed by the target and the distractor average. Finally, again according to the results ob-
56
57
58
59
60

361 tained by Fehd and Seiffert (2010), the gaze behavior should turn out to be independent of
 362 stimulus speed.

363 Results

364 As illustrated in Figure 2, results show significant response accuracy differences for
 365 the three speed conditions $F(2, 26) = 62.7, p < .01, \eta_p^2 = .83$, with the highest accuracies in
 366 the 6 °/s condition, followed by the 9 °/s and 12 °/s conditions, and each speed condition dif-
 367 fering in terms of response accuracies from each other (all $ps < .01$). The same data pattern is
 368 revealed for the average number of correctly identified targets $F(2, 26) = 54.8, p < .01, \eta_p^2 =$
 369 $.81$, all $ps < .01$, which means that the slower the objects moved, the more targets were cor-
 370 rectly recalled by the participants.



371

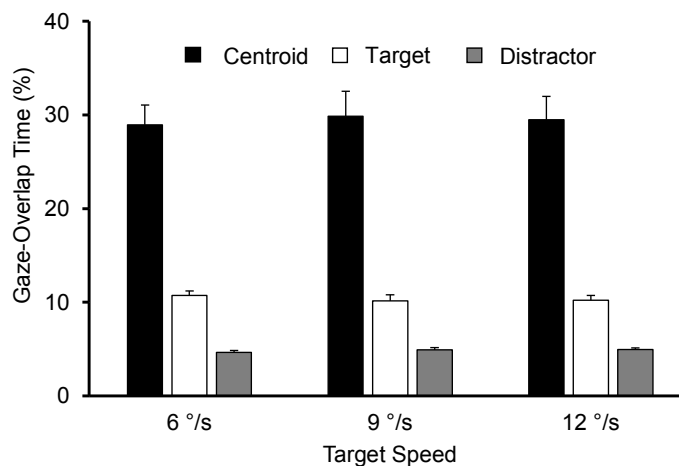
372 *Figure 2.* Response accuracy as percentage of correct responses (M and SE) and number of correctly recalled
 373 targets (M and SE) as a function of target speed.

374 In terms of gaze overlap to a priori-defined regions of interest, as depicted in Figure 3,
 375 a 3 (speed) x 3 (region) ANOVA with repeated measures on both factors revealed a signifi-
 376 cant main effect for region only, $F(2, 26) = 76.4, p < .01, \eta_p^2 = .86$. Gaze was allocated more
 377 to the centroid than to the targets or the distractors, while gaze overlapped more with targets

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

378 than with distractors, all p s < .01. Neither the effect for speed ($p = .90$, $1-\beta = .93$) nor the in-
 379 teraction of speed and region ($p = .92$, $1-\beta = .96$) reached significance.

380



381

382 *Figure 3.* Gaze-overlap time (M and SE) on the centroid, the target average, and the distractor average as a func-
 383 tion of target speed.

384 Discussion

385 Experiment 1 aimed to answer the question whether classical MOT findings on gaze
 386 patterns, in particular those reported by Fehd and Seiffert (2010), could be replicated in con-
 387 sideration of the novel experimental setting with a large projection and the introduction of a
 388 stationary-centroid phase. Regarding this replication trial, results show that the faster the ob-
 389 jects move, the more tracking accuracy is impaired (see Figure 2). This result is in perfect
 390 agreement with previous research (Alvarez & Franconeri, 2007; Fehd & Seiffert, 2010).
 391 Compared to the current findings, Fehd and Seiffert (2010) reported higher response accura-
 392 cies for the conditions 6 °/s (90 %) and 12 °/s (65 %), while the number of correctly tracked
 393 targets in our study is comparable with the numbers reported by Alvarez and Franconeri

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 394 (2007) at similar speed conditions (i.e., 7 °/s and 14 °/s) when spacing (1 stimulus diameter) is
4
5 395 alike (3.4 and 2.4 targets, respectively).
6
7

8 396 Regarding gaze behavior, Fehd and Seiffert (2010) found that participants were mainly
9
10 397 looking at the centroid of the target group, and this was true even at high object speeds. This
11
12 398 finding could also be replicated in the current study since results show about 29.5 % gaze
13
14 399 overlap with the centroid (Fehd & Seiffert, 2010: 25-30 %), 10.5 % with the target average
15
16 400 (Fehd & Seiffert, 2010: approx. 10 %), and 4.8 % with the distractor average (Fehd & Seif-
17
18 401 fert, 2010: approx.. 5 %; see Figure 1B in Fehd & Seiffert, 2010). A further analysis of gaze
19
20 402 switches revealed that center-target switches ($M = 5.5$, $SD = 1.9$) were more often used than
21
22 403 target-target switches ($M = 1.3$, $SD = 0.6$) which again is in perfect line with previous results
23
24 404 (Fehd & Seiffert, 2010: Exp 1: 5.9 vs. 1.9). Thus, it could be demonstrated again that the cen-
25
26 405 troid seems to have a pronounced value in MOT. From a vision perspective, an obvious ex-
27
28 406 planation for this gaze behavior is that peripheral information about target positions are used
29
30 407 to effectively monitor all targets. On the basis of this finding, the central research question on
31
32 408 the role of peripheral vision in monitoring moving objects regarding unexpected changes can
33
34 409 be reasonably asked.
35
36
37
38
39
40
41
42

411 **Experiment 2**

412 In Experiment 2, the functionality of peripheral vision in tracking multiple objects will
43
44 413 be under investigation. In particular, the study aims to answer the question whether partici-
45
46 414 pants are able to respond to a target-related change in MOT peripherally, that means, before
47
48 415 gaze is on this target. Furthermore, since motion information are used for tracking and be-
49
50 416 cause of a high temporal resolution of the peripheral retina, it is also predicted that motion
51
52 417 changes are detected better than form changes if both kinds of change detections feature a
53
54
55
56
57
58
59
60

1
2
3 418 comparable degree of task difficulty. For analyzing events such as motion or form changes,
4
5 419 the evaluated stationary-centroid phase from Experiment 1 will be important, as this manipu-
6
7 420 lation opens the door for precisely controlling eccentricities of events. For evaluating the dy-
8
9 421 namics of the visual search behavior that underlies change identification, gaze will be ana-
10
11 422 lyzed using a saccade-detection algorithm to determine the “saccadic reaction time”, which is
12
13 423 defined as the time interval between the onset of a change and the onset of the saccade. It is
14
15 424 expected to find differences in saccadic reaction times between the two manipulation condi-
16
17 425 tions because if both kinds of change detections are of comparable task difficulty, due to mo-
18
19 426 tion sensitivity of peripheral vision, a motion change should be better detected with peripheral
20
21 427 vision than form change. More specifically, as peripheral vision is often used to guide sac-
22
23 428 cades (Knowler, 2011) and a rescue saccade is more likely to be used in the form-change con-
24
25 429 dition to perceive the change with overt attention (i.e., while the square is still a diamond),
26
27 430 shorter saccadic reaction times are expected in this case, which would also indicate that par-
28
29 431 ticipants rely less on peripheral vision in the form-change condition.
30
31
32
33

34 432 In Experiment 2, the motion change was implemented as a sudden target stop from a
35
36 433 speed of $6^\circ/\text{s}$ and the form change as a square-to-diamond change over the stationary-
37
38 434 centroid phase. These specifications were foremost made for reasons of ecological validity,
39
40 435 which means, in reference to change-detection demands in natural settings. When illustrating
41
42 436 this point by the sports examples sketched in the beginning of this paper, in a tennis match
43
44 437 with an opponent moving close to the court’s opposite baseline, this motion change would
45
46 438 correspond to an opponent’s sudden stop of a – rather slow – running movement at a speed of
47
48 439 about 2.5 m/s whereas the form change would approximately correspond to an opponent’s
49
50 440 racket movement of 30 cm to the left or right side of his or her body. Hence, from an applied
51
52 441 perspective, the experimental conditions can definitely claim ecological validity which is im-
53
54
55
56
57
58
59
60

1
2
3 442 portant when it comes to the derivation of recommendations for real-life practice like sports
4
5 443 training.
6
7

8 444 However, based on these specifications and with regard to the above-made predictions
9
10 445 on superior detections rates for motion changes, the only point that could be inferred from an
11
12 446 empirical confirmation would be that motion changes *as they appear in natural settings* can
13
14 447 be better detected than form changes *as they appear in natural settings*. From an experimental
15
16 448 point of view, this inference satisfies to a limited degree only, as the alternative explanation
17
18 449 that superior motion-change detection rates are mainly caused by higher salience cannot be
19
20 450 ruled out. The crucial argument would then be that the specified motion change simply fea-
21
22 451 tures a minor task difficulty which could be turned around either by a difficulty increase of
23
24 452 the motion-change detection task (e.g., slowing down the target speed from 6 °/s to 5 °/s) or
25
26 453 by a difficulty decrease of the form-change detection task (e.g., if the target becomes a cross
27
28 454 of twice the original size). At this point, an objective measure of the respective task difficulty
29
30 455 would be desirable in order to be able to choose motion and form changes of a comparable
31
32 456 difficulty. However, due to the dimensional character of the two types of changes, such a
33
34 457 simple measure is not available for a principal reason explained by Rensink (2002) in his re-
35
36 458 view on change-detection research as follows: “Performance depends on the magnitude of the
37
38 459 change and there is no simple way to equate the visibility of different kinds of changes”
39
40 460 (p. 255).
41
42
43
44
45

46 461 For this reason, in order to empirically control for task-difficulty differences between
47
48 462 the two change conditions, additional data was gathered by putting the two types of changes
49
50 463 into foveal vision and collecting response times (as detection rates are generally 100 % for
51
52 464 foveal vision). Response times of the same magnitude could then be taken as a hint that the
53
54 465 specified changes do *not* differ in task difficulty per se such that a potentially revealed superi-
55
56 466 ority of motion-change detection rates by peripheral vision should be foremost ascribed to
57
58
59
60

1
2
3 467 specifics of the peripheral visual system. In the following, details of this task-difficulty check
4
5 468 (TDC) will be reported at the end of each Methods subsection as well as at the beginning of
6
7 469 the Results section.
8
9

10 470

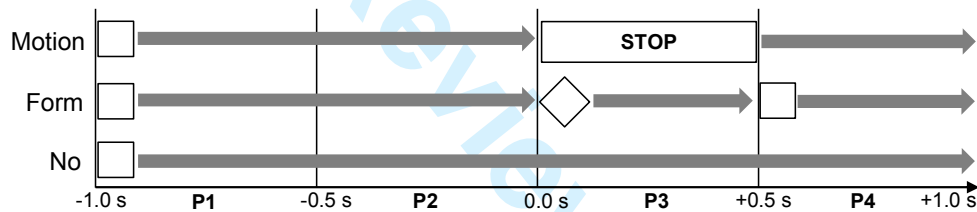
11 12 13 471 **Methods**

14
15
16 472 *Participants.* Since performing a MOT task and additionally detecting target changes
17
18 473 is a quite complex experimental task, it seemed advisable to conduct the study with partici-
19
20 474 pants who were familiar with the basic task of MOT. For this reason, Experiment 2 was run
21
22 475 with the participants who had already taken part in Experiment 1 ($N = 14$), accounting the 135
23
24 476 trials of Experiment 1 as familiarization phase for the increasing task demands. Experiment 2
25
26 477 was conducted five weeks after the finalization of Experiment 1. For the additional TDC
27
28 478 study, 36 volunteers were recruited (12 females and 24 males; age: 29.25 ± 6.65 years). Both
29
30 479 experiments were undertaken in accordance with the Declaration of Helsinki.
31
32

33
34 480 *Stimuli.* Stimuli were created in the same way as in Experiment 1 but differed regard-
35
36 481 ing speed, which was set to $6^\circ/\text{s}$ for all trials since response accuracy at this speed was high-
37
38 482 est in Experiment 1. Furthermore, motion and form changes of one of the targets were intro-
39
40 483 duced. These changes were evoked in constant relation to the onset of the stationary-centroid
41
42 484 phase, which, as in Experiment 1, started at either 3.0 s, 3.5 s, or 4.0 s after motion onset of
43
44 485 the objects. Unknown to the participants, the change always occurred exactly 250 ms after the
45
46 486 collision of the target with the distractor, which caused the stationary-centroid phase. At this
47
48 487 point of time, depending on the trial's specific manipulation condition, one out of the four
49
50 488 targets either stopped for 0.5 s before continuing the motion (motion change) or abruptly be-
51
52 489 came a diamond (i.e., the square was rotated by 45°) for 0.5 s but without altering the current
53
54 490 speed (form change) (see Figure 4). Furthermore, to ensure a pre-defined eccentricity of
55
56
57
58
59
60

1
2
3 491 events, the distance of the event-target was forced to be 15° (if participants, as expectable
4
5 492 from Experiment 1, actually pursued a centroid strategy). To prevent participants from antici-
6
7 493 pation of the changes, a control condition without any target changes was included. As each
8
9 494 control trial matched a change trial with a particular motion onset, despite the absence of any
10
11 495 change, the control trials were assigned to a (virtual) stationary-centroid phase onset condition
12
13 496 in accordance to this match.

14
15
16
17 497 The same stimulus configuration was used in the TDC experiment. This time, howev-
18
19 498 er, at the beginning of a single trial, only one target was highlighted with a red frame, namely
20
21 499 the target which would (virtually) change its motion or form over the subsequent motion
22
23 500 phase. Participants were instructed to keep this target in foveal vision during the whole trial.



501

502 *Figure 4.* Temporal properties of the three target-change conditions. In phase P3, the target either stopped (Mo-
503 tion) or became a diamond (Form) or continued to move unchanged (No) for 0.5 s whereas no differences be-
504 tween the conditions are evident for the remaining phases P1, P2, and P4.

505 *Apparatus.* In addition to the setup used in Experiment 1, a single-button response de-
506 vice (1000 Hz) was integrated into to the VICON data collection system to measure motor
507 response times.

508 *Procedure.* Experiment 2 was organized in individual single sessions that lasted about
509 one hour. Participants' task was to press the button as soon as a motion- or form-related target
510 change occurred and to name the number of the changed target at the end of the trial. Howev-
511 er, if no change occurred, the task was to recall all four targets. Participants initially practiced
512 each of the three conditions with three trials that were not shown later in the test blocks, re-

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 513 sulting in 9 practice trials in total. Taking the stimulus constraints into account, a total of 120
4
5 514 trials were derived from 12 mother trials. These trials were quasi-randomly ordered in 10
6
7 515 blocks, with 12 trials each with the additional specification that per block each mother trial
8
9 516 appeared exactly once and each change/control condition and each onset of the stationary-
10
11 517 centroid phase exactly four times. As in Experiment 1, verbal decisions were recorded in writ-
12
13 518 ing, and no knowledge of the results was given after the test trials.

14
15
16
17 519 In the TDC experiment, participants started with 6 familiarization trials followed by
18
19 520 30 test trials in 2 blocks (10 trials for form, motion and no-change condition each) in which
20
21 521 they had to press the button as soon as a motion- or form-related target change occurred. In
22
23 522 addition, at the end of each trial, participants had to name the number of the target they had
24
25 523 followed with their eyes.

26
27
28 524 *Gaze analysis.* Since in Experiment 2 the focus was laid on the dynamics of gaze be-
29
30 525 havior as a function of monitoring object changes with peripheral vision, two kinds of trials
31
32 526 had to be excluded from further analysis. First, a trial could not be included if the target
33
34 527 change was not correctly signaled by a button press or not correctly named at the end of the
35
36 528 trial's motion phase, as in those trials attention was not on the correct target. Second, trials
37
38 529 were excluded from further analysis if, at event start, the distance of the gaze vector to the to-
39
40 530 be-changed target was less than 5° , that is, within the range of parafoveal vision (Calvo &
41
42 531 Lang, 2005). The reason for this exclusion is that in those trials the participants could have
43
44 532 focused on the respective target just by chance so that the missing need for a saccade should
45
46 533 be considered as accidental and not be interpreted as resulting from a certain gaze-control
47
48 534 strategy. In the Results section, detailed information will be given on the percentage of trials
49
50 535 that had been excluded due to these criteria.
51
52
53
54
55
56
57
58
59
60

1
2
3 536 For the remaining trials, two gaze-related groups of variables were aggregated. First,
4
5 537 the distance of the gaze vector to the target as well as to the centroid was calculated (in °) as
6
7 538 the average value for four phases P1-P4 of 0.5 s duration each. As depicted in Figure 4, these
8
9 539 phases were defined in relation to the change event (in case of control trials, to the virtual
10
11 540 change event) with P1 beginning 1.0 s before, P2 beginning 0.5 s before, P3 beginning exact-
12
13 541 ly with the target change and P4 beginning exactly with the change's termination. After the
14
15 542 calculation of the mean gaze-target and gaze-centroid distances for the four phases per single
16
17 543 trial, the values were averaged for each participant over the valid trials per change/control
18
19 544 condition.

20
21
22
23 545 To analyze the dynamics of the peripheral monitoring of the targets, for the valid mo-
24
25 546 tion- and form-change trials, a second group of gaze-related variables referred to the onset and
26
27 547 offset of the first saccade to the changed target. For this purpose, saccades were identified by
28
29 548 a velocity-based detection algorithm with adaptive thresholds on the basis of local noise lev-
30
31 549 els (Nyström & Holmqvist, 2010). This analysis resulted in three time intervals (in ms), which
32
33 550 (in cases in which the identification was signaled by the participants after the saccade) accu-
34
35 551 mulated to the motor-response time indicated by the button press: a saccadic reaction time
36
37 552 from the onset of the target change to the onset of the saccade, a saccadic duration from the
38
39 553 onset to the offset of the saccade, and a fixation duration from the offset of the saccade to the
40
41 554 button press. For each participant, the values per trial were averaged over the valid trials for
42
43 555 motion- and form-related changes. It can be argued that the shorter the saccadic reaction is,
44
45 556 the less is the change detection based on peripheral vision.

46
47
48
49
50
51 557 In the TDC experiment, it was controlled that participants used foveal vision for de-
52
53 558 tecting target changes (maximum distance of 3 ° of visual angle to the target).
54
55
56
57
58
59
60

1
2
3 559 *Measures.* Besides the gaze-related variables, response accuracy (in %) was gathered,
4
5 560 in the no-change trials as the percentage of trials in which all four targets could be correctly
6
7 561 recalled, and in the motion- and form-change trials as the percentage of trials in which the
8
9 562 target change was correctly signaled by the button press and the correct number of the event-
10
11 563 target was named at the end of the trial as well. Furthermore, motor-response time was calcu-
12
13 564 lated (in ms) for all correct motion- and form-change trials as the time between target-change
14
15 565 onset and button-press and then averaged over all available trials. Response time calculations
16
17 566 in the TDC experiment were conducted in the same manner.

18
19
20
21 567 All dependent variables were analyzed with repeated-measures ANOVAs, including
22
23 568 either all three or just the two target-change conditions as a first factor and, if applicable, a
24
25 569 gaze- or phase-related variable as a second or third factor. Significant main or interaction ef-
26
27 570 fects were further analyzed with paired t-tests. The alpha level for tests on differences was a
28
29 571 priori set to $\alpha = .05$, a posteriori effect sizes were computed as partial eta squares (η_p^2), and
30
31 572 the power ($1-\beta$) was calculated in cases of non-significant tests.

32
33
34
35 573 It was expected to find higher response accuracies in the motion- compared to the
36
37 574 form-change condition, while response accuracies in the no-change condition should not dif-
38
39 575 fer from the results obtained in Experiment 1 for the same stimulus speed. Furthermore, it was
40
41 576 hypothesized that the average gaze distance to the centroid is smaller than to the manipulated
42
43 577 target in P1 and P2. However, gaze should be closer to the target in the form- and motion-
44
45 578 change condition in P4 (latest), whereas in the no-change condition, the gaze distance should
46
47 579 be continuously smaller to the centroid than to the target. For the TDC experiment, it was
48
49 580 predicted that the participants were *not* able to respond faster to foveally perceived motion
50
51 581 changes than to foveally perceived form changes. Under the assumption that this prediction
52
53 582 was confirmed, a superior performance for motion than for form changes was anticipated for
54
55 583 the peripheral viewing conditions of the main experiment. With respect to the dependent vari-
56
57
58
59
60

1
2
3 584 ables described above, it was particularly expected that the motor response can be executed
4
5 585 ahead of the saccade onset in more trials for the motion-change than in the form-change con-
6
7 586 dition.
8
9

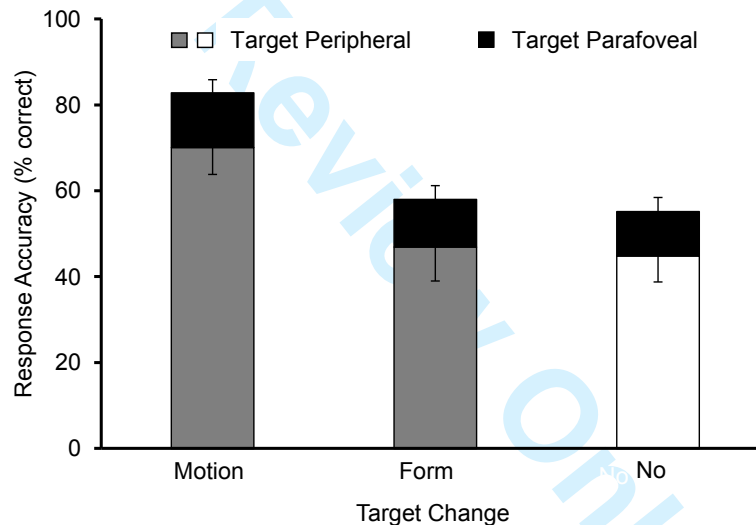
10 587

11 12 13 588 **Results**

14
15
16 589 For the TDC experiment, paired t-tests were used to test response time differences for
17
18 590 motion vs. form change detections and Cohen's d was calculated as effect size. Data of 35
19
20 591 participants could be analyzed. Results not only revealed that motion changes were *not* de-
21
22 592 tected faster than for form changes, as it has been predicted above, but that form changes were
23
24 593 actually detected significantly faster than motion changes when being put in foveal vision
25
26 594 (motion: $M = 427.8$ ms, $SE = 72.3$ ms; form: $M = 289.1$ ms, $SE = 48.9$ ms), $t(34) = 11.39$, $p <$
27
28 595 $.01$, $d = 1.7$.

29
30
31
32 596 For the main experiment, three participants had to be excluded from the analysis be-
33
34 597 cause they were only able to perceive a target change if central gaze was by chance on the
35
36 598 respective target leading to less than five trials that fulfilled the above-defined inclusion crite-
37
38 599 ria. On the contrary, the remaining 11 participants clearly pursued a centroid-tracking strate-
39
40 600 gy, as at target-change onset, the respective target was out of parafoveal vision in 83.2 % (SE
41
42 601 = 4.1 %) of the correctly identified motion-change and in 78.0 % ($SE = 6.4$ %) of the correctly
43
44 602 identified form-change trials, respectively. When applying the same calculation to the corre-
45
46 603 sponding no-change conditions, the target that would change in the change-derivatives of the
47
48 604 same mother trial, one receives an average of 79.4 % ($SE = 6.4$ %) of trials in which all the
49
50 605 four targets could be successfully recalled and the respective target was out of parafoveal vi-
51
52 606 sion at the moment of the (virtual) target change. To provide consistency, the following re-
53
54 607 sults refer to the 11 selected participants and to the just-described percentage of valid trials.
55
56
57
58
59
60

As depicted in Figure 5, participants were able to detect target-motion changes better than form changes. This is confirmed when the ANOVA is based on all trials in which the change could be identified correctly (grey plus black fractions of the bars), $F(1, 10) = 17.2, p < .01, \eta_p^2 = .63$, as well as when the calculation is restricted to the trials that were valid in that sense that, at target-change onset, the respective target was out of parafoveal vision (grey fractions of the bars), $F(1, 10) = 10.71, p < .01, \eta_p^2 = .52$. Furthermore, response accuracy in the no-change condition ($M = 55.2\%$, $SE = 4.5\%$; white plus black fraction of the bar) showed no significant differences to those obtained in Experiment 1 for the same object speed ($M = 59.6\%$, $SE = 6.1\%$), $t(11) = 0.82, p = .43, 1-\beta = .74$.



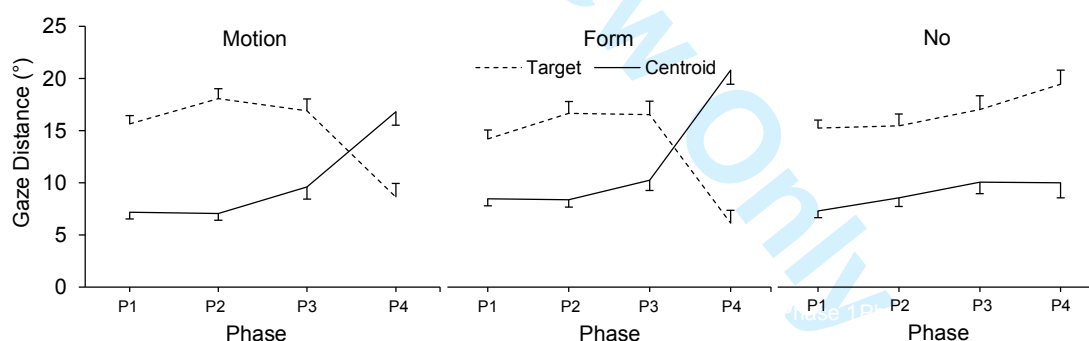
617

Figure 5. Response accuracy (M and SE) as a function of target-change condition. In the motion- and form-change conditions, a trial was counted as correct if the button was pressed and the correct target number was named at the end of the trial whereas, in the no-change condition, all four targets had to be recalled correctly. The black fractions of the bars depict the percentage of trials that were excluded from further analyses because, at change onset, the respective target was already in parafoveal vision.

For the event-related gaze-distance analysis, which is illustrated in Figure 6, a 3 (target change) \times 2 (location) \times 4 (phase) ANOVA with repeated measures on all three factors was conducted. The results show significant main effects for location, $F(1, 10) = 8.21, p = .02, \eta_p^2$

625

626 = .45, and phase, $F(3, 30) = 30.7, p < .01, \eta_p^2 = .76$, as well as the following significant two-
 627 way interactions: target change x location, $F(2, 20) = 21.3, p < .01, \eta_p^2 = .68$, target change x
 628 phase, $F(6, 60) = 4.9, p < .01, \eta_p^2 = .33$, and location x phase, $F(3, 30) = 42.6, p < .01, \eta_p^2 =$
 629 .81. Furthermore, a significant three-way interaction target change x location x phase was
 630 found, $F(6, 60) = 83.8, p < .01, \eta_p^2 = .89$. Consequently, to further interpret the three-way
 631 interaction, paired t-tests were used to evaluate gaze-distance differences to the centroid com-
 632 pared to the manipulated target for all target-change conditions and phases. The results show
 633 that, for all target-change conditions, the gaze vector was closer to the centroid than to the
 634 respective target in the first three phases (all $ps \leq .01$), while in P4, gaze was closer to the
 635 target in the motion-change condition as well as in the form-change condition (all $ps < .01$),
 636 whereas the opposite was true in the no-change condition ($p < .01$). Furthermore, the gaze
 637 vector was closer to the target in P4 of the form trials compared with P4 of the motion-change
 638 trials ($p = .03$).



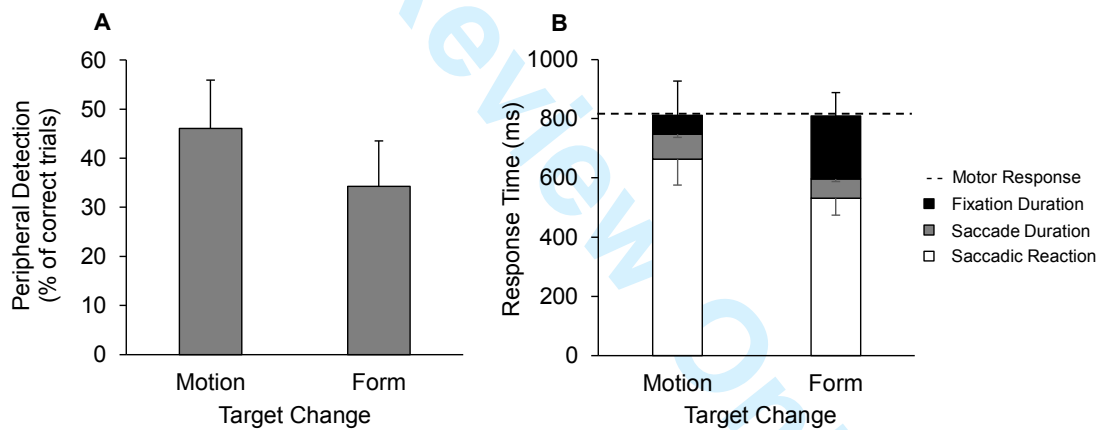
639

640 *Figure 6.* Gaze distance (M and SE) to the centroid and the changing target for four phases (P1-P4), separately
 641 depicted for three target-change conditions. In the no-change condition, gaze distances were calculated to the
 642 target that changed in the matched trials of the other conditions.

643 In Figure 7, the dynamics of the interplay between peripheral monitoring and saccadic
 644 reaction from the centroid to the changed target is illustrated in more details. As depicted in
 645 Figure 7A, the button was pressed before gaze was on the target in a higher percentage of

60

646 correct trials in the motion-change condition compared to the form-change condition, $F(1, 10)$
 647 $= 8.9$, $p = .01$, $\eta_p^2 = .47$, which means that in those cases, the change was definitely detected
 648 by means of peripheral vision. Furthermore, as shown in Figure 7B, a respective ANOVA
 649 reveals significant differences in saccadic reaction time after target-change onset, $F(1, 10) =$
 650 6.7 , $p = .03$, $\eta_p^2 = .40$, as well as in fixation time before the motor response was indicated by
 651 the button press, $F(1, 10) = 7.1$, $p = .02$, $\eta_p^2 = .42$, whereas no differences were found for sac-
 652 cade duration, $F(1, 10) = 1.6$, $p = .24$, $1-\beta = .94$. This means that later saccade onsets to the
 653 changed target were observed in the motion-change condition compared to the form-change
 654 condition, while in the form-change condition, targets were fixated longer before the motor
 655 response was executed.



656
 657 *Figure 7. (A) Peripheral change detection (M and SE) as a function of target change, calculated as the percentage*
 658 *of trials in which the motor response was executed before the gaze vector reached the respective target. (B) Sac-*
 659 *cadic reaction, saccade duration, and fixation duration (M and SE) as cumulated response up to the average mo-*
 660 *tor response time. The value “0” on the vertical axis signifies the onset of the target change.*

661

662 Discussion

663 In Experiment 2, a novel approach in studying MOT performance has been introduced
 664 by using an event-driven saccade-based gaze analysis that precisely allows for examining

1
2
3 665 whether motor responses can be executed when changes in the environment are perceived
4
5 666 with peripheral vision only. First of all, for trials without a change, intra-individual response
6
7 667 accuracy did not differ from the values obtained in Experiment 1, which could be seen as an
8
9 668 indicator that the additional detection task did not affect tracking performance negatively (as
10
11 669 it was also suggested by Bahrami, 2003). Furthermore, no learning effects were observed over
12
13 670 the five weeks between both experiments. On this basis, with respect to the main research
14
15 671 question, it could be clearly shown that motor responses can be reliably initiated if the deci-
16
17 672 sive stimuli can be perceived by peripheral vision only. Moreover, as hypothesized, higher
18
19 673 response accuracies for the motion-change condition could be revealed than for the form-
20
21 674 change condition. For the appropriate interpretation of this finding the results of the TDC
22
23 675 study must be taken into account as they show faster responses to form than to motion chang-
24
25 676 es (of exactly the same magnitude as in the main experiment) when the changes are put into
26
27 677 foveal vision such that the superior detection of motion changes in the main experiment can
28
29 678 definitely not be ascribed to a mere task-difficulty effect but must be attributed to specifics of
30
31 679 the peripheral visual system.

32
33
34
35
36
37 680 Having a closer look onto the main results, the gaze distance was significantly closer
38
39 681 to the centroid than to the targets over the two phases before the event (P1 and P2) and even
40
41 682 over the phase of the change (P3), whereas over the last phase (P4), gaze was mostly on the
42
43 683 target in change trials whereas it stayed closer to the centroid in no-change trials. At this
44
45 684 point, it should be noted that the gaze distances depicted in Figure 6 resulted from an averag-
46
47 685 ing procedure which means that a relatively small percentage of trials with no saccades to the
48
49 686 target result in large gaze-target distances in P4 of the change conditions. The same is true for
50
51 687 the gaze-centroid distances in P1-P3 and, in addition, in P4 of the no-change condition since
52
53 688 saccades to targets are occasionally initiated (Elfanagely et al., 2011; Fehd & Seiffert, 2008,
54
55 689 2010), thereby increasing the mean distance to the centroid. For this reason, the respective
56
57
58
59
60

1
2
3 690 values ranging between 5° and 10° should, in the end, be understood as a result of focusing
4
5 691 either a target or the centroid in most of the cases, but not always. Hence, the data should be
6
7 692 interpreted in such a way that, in cases of a correct response, participants produce a saccade
8
9 693 from the centroid to the respective target whenever a target changes its motion or form.

10
11
12 694 However, when bearing in mind that the target change over P3 lasted 0.5 s only and
13
14 695 that the change was terminated at the beginning of P4, this finding should not be misinterpret-
15
16 696 ed in such a way that the saccades revealed for P4 are mandatorily needed to detect the re-
17
18 697 spective change. This conclusion is underpinned by the gaze data depicted in Figure 7, show-
19
20 698 ing that mean saccadic-reaction times are longer than 0.5 s in both target-change conditions
21
22 699 and that, in a considerable percentage of trials, the saccade was initiated even after the button
23
24 700 press, that is, on average, more than 0.8 s after the target change. Hence, in a lot of cases, the
25
26 701 decision must have been made before a verification by a suitable fixation could have taken
27
28 702 place.

29
30
31
32
33 703 When comparing the two target-change conditions in more details, in line with our
34
35 704 predictions, in significantly more correct trials in the motion-change compared with the form-
36
37 705 change condition, the button press had been executed before the saccade to the changed target
38
39 706 was initiated. In both cases, as already argued above, these changes must have been perceived
40
41 707 with peripheral vision since the gaze position was more than 5° away from the changing tar-
42
43 708 get. Furthermore, saccadic reaction time turned out to be longer in the motion-change com-
44
45 709 pared with the form-change condition supporting the assumption that foveal information are
46
47 710 not as much required for motion-change as for form-change detection. On the contrary, as
48
49 711 hypothesized, a time of gaze being on the changed target seems to be required in the form-
50
51 712 change condition before the button press can be executed (which would also be in line with
52
53 713 the results of the TDC experiment that form changes can be detected faster with foveal vi-
54
55 714 sion).

1
2
3 715 To test whether this saccade is helpful to respond faster, a “saccading strategy” (i.e.,
4
5 716 gaze is on the target *before* the button is pressed) and a “peripheral strategy” (i.e., gaze is on
6
7 717 the target *after* the button is pressed) can be distinguished and checked against each other by
8
9 718 post-hoc paired t-tests for both manipulation conditions and gaze strategies. Eight participants
10
11 719 could be integrated into this analysis because they exhibited correct responses with both gaze
12
13 720 strategies in both conditions. Results show that response times for the two gaze strategies nei-
14
15 721 ther differ in the form- ($p = .93$) nor in the motion-change condition ($p = .40$), indicating that
16
17 722 initiating a saccade to the target does *not* help to respond faster compared to using peripheral
18
19 723 vision only. Thus, initiating a saccade to the target does *not* lead to faster responses but could
20
21 724 be interpreted as just being helpful to early track the target for naming the correct number at
22
23 725 the end of the trial.
24
25
26
27
28
29
30

31 727 **Overall Discussion**

32
33
34 728 The current study aimed to investigate whether peripheral vision can be used to moni-
35
36 729 tor multiple moving objects as well as to detect single-target changes. For this purpose, the
37
38 730 designated experimental MOT setup with a large projection and a constant-position centroid
39
40 731 phase had to be checked first. Therefore, in Experiment 1, the successful replication of the
41
42 732 findings regarding the use of a virtual centroid to track multiple objects (c.f., Fehd & Seiffert,
43
44 733 2008, 2010) and the dependency of tracking accuracy on target speed (c.f., Cavanagh & Alva-
45
46 734 rez, 2005) had been important for conducting the main study. In the subsequent Experiment 2,
47
48 735 it could be clearly shown that peripheral vision is used at great extents to detect changes in
49
50 736 MOT. Moreover, for the change conditions specified on the basis of considerations regarding
51
52 737 the ecological validity of the manipulations, it could be demonstrated that motion changes can
53
54
55 738 be better detected peripherally than form changes. Finally, an additionally conducted control
56
57
58
59
60

1
2
3 739 experiment showed that this superiority must be explained by particularities of the peripheral
4
5 740 visual system and cannot be attributed to possible differences in task difficulty.
6
7

8 741 As debated in the introductory section, findings on MOT are often discussed exclu-
9
10 742 sively from an attention perspective. However, based on the gaze-related findings revealed in
11
12 743 the experiments at hand, it seems worthwhile to consider classical MOT results also from a
13
14 744 vision perspective. In this regard, peripheral vision (i.e., perceiving stimuli with peripheral
15
16 745 vision) and spatial covert attention (i.e., allocating attention without eye movements; Stras-
17
18 746 burger et al., 2011) seem to describe closely related phenomena. Therefore, in the following
19
20 747 paragraphs, five reasons will be discussed regarding an integration of properties of the visual
21
22 748 system in the context of attention-related findings to understand the role of covert attention in
23
24 749 MOT. These reasons refer to (a) the ecological validity of methods to measure covert atten-
25
26 750 tion, (b) the detection of feature changes that depend on visual capabilities, (c) the connection
27
28 751 of covert attention to a centroid-MOT strategy and consequences of saccade-related costs for
29
30 752 processes of attentional monitoring, (d) the fact that eye-tracking devices can be generally
31
32 753 useful for testing assumptions made in attentional theories on MOT, and (e) spatial uncertain-
33
34 754 ty in the context of covert-attention processes, which could also be related to the voluntary
35
36 755 control of saccades. At the end of each paragraph, a follow-up question for future experiments
37
38 756 will be derived.
39
40
41
42
43

44 757 (a) In classical MOT studies, participants were usually instructed to fixate a central
45
46 758 point to examine covert attention in detecting feature changes (e.g., Sears & Pylyshyn, 2000).
47
48 759 This approach, however, seems problematic since, first, the fixation instruction was not
49
50 760 checked by the application of gaze-registration methods and, second, fixating a central point
51
52 761 may not reflect the participant's natural behavior. In contrast, in the current experiment, par-
53
54 762 ticipants were tested under ecologically valid conditions, being free to move their eyes and
55
56 763 their head because the actual gaze could be controlled by eye- and head-tracking devices. Be-
57
58
59
60

1
2
3 764 yond these advantages regarding ecological validity, the event-driven saccade-based gaze
4
5 765 analysis offers the opportunity to estimate whether peripheral or foveal vision was used to
6
7 766 detect changes leading to enriched inferences on the role of covert attention in MOT. Thus, in
8
9
10 767 future research, by applying eye-tracking methods, covert-attention processes may be exam-
11
12 768 ined even in real-life settings in which free gaze behavior is allowed.

13
14
15 769 (b) The use of covert attention not only depends on psychological but also on visual
16
17 770 capabilities, especially when target changes have to be detected in MOT. In former studies
18
19 771 (e.g., Bahrami, 2003) participants regularly showed better detection rates for color changes
20
21 772 compared with form changes, but, as a matter of the respectively pursued research questions,
22
23 773 visual capabilities were not taken into account. Thus, it is not known so far whether the supe-
24
25 774 riority regarding color-change detection actually results from specifics of attentional process
26
27 775 or could be attributed to particularities of the visual system. In the current study, those visual
28
29 776 capabilities were taken into account because finding better detection rates of motion com-
30
31 777 pared with form changes was expected due to the light and motion sensitivity of the peripher-
32
33 778 al human retina, which comes along with a higher sensitivity to detect motion changes (Fin-
34
35 779 lay, 1982; Goldstein, 2010; McKee & Nakayama, 1984). Indeed, the frequency of detection
36
37 780 was found to be higher in the motion-change condition; however, the contrast sensitivity of
38
39 781 the peripheral retina obviously allowed detecting form changes, too. Consequently, the lower
40
41 782 detection rates for form changes can be traced back to the lower spatial resolution of the pe-
42
43 783 ripheral visual system. Hence, in a follow-up experiment, the low spatial acuity but high mo-
44
45 784 tion sensitivity of the peripheral retina should be further investigated by manipulating the ec-
46
47 785 centricities of changes. In the case of a visual-system effect on detection rates, the prediction
48
49 786 can be derived that, as a function of eccentricity, form-change detection should be impaired to
50
51 787 a higher degree than motion-change detection.
52
53
54
55
56
57
58
59
60

1
2
3 788 (c) Previous research has shown that covert attention is used for tracking purposes
4
5 789 (Fehd & Seiffert, 2008, 2010). The gaze-related findings of the current study extend this view
6
7 790 by showing that a centroid strategy is preferred not only for tracking, but also for change de-
8
9 791 tection. This gaze strategy is presumably favored because a target-switching strategy with an
10
11 792 alternating gaze between targets (Fehd & Seiffert, 2010) would cause a lot of costs due to the
12
13 793 suppression of information processing before, during, and after the saccade (Diamond, Ross,
14
15 & Morrone, 2000). As a consequence, the chances to miss the target-feature change would be
16
17 794 higher. Participants in our study indeed seem to avoid costly eye movements by using the
18
19 795 virtual centroid strategy, in attentional terms: by monitoring the targets with covert attention
20
21 796 and “waiting” for a feature change. With the results at hand it can be shown that covert atten-
22
23 797 tion is not only used to monitor the distances between targets and barrier occluders (Zelinsky
24
25 798 & Todor, 2010) but also to detect target changes. In future experiments, this hypothetical
26
27 799 mechanism could be investigated further by manipulating the costs of eye movements. In
28
29 800 more details, due to the suppression of information processing as a consequence of saccades,
30
31 801 it would be expected that the more saccades are experimentally induced, the more likely a
32
33 802 change will remain undetected.
34
35 803
36
37
38

39 804 (d) By studying natural gaze behavior, assumptions made by theoretical models on the
40
41 805 role of attention in MOT can be assessed more rigidly. On the one hand, our findings confirm
42
43 806 attentional theories in different respects. In this regard, the observed gaze behavior indicates
44
45 807 the value of the center of the polygon for MOT performance (Fehd & Seiffert, 2010), which is
46
47 808 in line with the model proposed by Yantis (1992). Furthermore, the higher detection rates in
48
49 809 the motion-change condition can be explained by the fact that the anticipated target locations
50
51 810 of the polygon (Iordanescu, Grabowecky, & Suzuki, 2009) are no longer perceptible because
52
53 811 of a target stop, whereas the polygon still exists in the form-change condition causing lower
54
55 812 detection rates. In this case, presumably because of the high tracking load with four targets
56
57
58
59
60

1
2
3 813 (Fencsik et al., 2007, Howe & Holcombe, 2012; Luu & Howe, 2015), no motion extrapolation
4
5 814 seems to be used in MOT. Hence, change detection results could be better explained with the
6
7 815 multifocal theory and concurrent streams on the targets where position information are used
8
9 816 for tracking as participants were able to detect motion as well as form changes. Thus, a posi-
10
11 817 tion tracker and a stream of object features seem to be concurrently used to keep track of the
12
13 818 targets. On the other hand, however, our gaze analyses also pose challenges for attentional
14
15 819 MOT theories. For example, in the model proposed by Cavanagh and Alvarez (2005), the
16
17 820 relative impact of the motion tracker and feature stream has to be questioned, as it can be in-
18
19 821 ferred from the results at hand that, when using covert attention to keep track of the targets,
20
21 822 the position tracker has a more pronounced sensitivity in the concurrent-tracking procedure
22
23 823 because motion changes were detected more often than form changes. This line of thought
24
25 824 would be supported by the weighting-mechanism proposed by Papenmeier et al. (2014) be-
26
27 825 cause spatiotemporal information during tracking are reliable until a motion change is induced
28
29 826 but a “false” spatiotemporal reliability is given in the form change condition. Hence, the de-
30
31 827 crease in distinctiveness of spatiotemporal and feature information could explain lower detec-
32
33 828 tion rates in the form change condition. Besides the relation of both streams, the concurrency
34
35 829 of attentional streams was challenged with our results because a number of changes were not
36
37 830 detected. Thus, it could be that streams are not permanently existent. Instead, covert attention
38
39 831 seems to switch between the targets leading to a limited amount of time to process target in-
40
41 832 formation which, in turn, would increase the difficulty to detect a change and explain missed
42
43 833 changes. Another challenge arises in the context of the FINST-model, in which it is assumed
44
45 834 that focal attention must be directed to the targets to update their positions (Fehd & Seiffert,
46
47 835 2010; Oksama & Hyönä, 2004). Contrary to this assumption, our research shows that periph-
48
49 836 eral vision, that is, covert attention, can be used to update target features and to detect feature
50
51 837 changes whereas foveal vision, that is, focal attention, seems mainly to be used to keep track
52
53
54
55
56
57
58
59
60

1
2
3 838 of the changed target to correctly name its number at the end of the trial. In particular, the
4
5 839 assumption of the feature-blind visual index model by Pylyshyn (1989, 2007), that feature
6
7 840 information cannot be accessed during tracking, is challenged by our data. On the contrary,
8
9 841 our results support the findings by Papenmeier et al. (2014), that spatiotemporal *and* feature
10
11 842 information can be processed. Consequently, future research on MOT should be directed to
12
13 843 the empirical test of specific predictions of attention-related models by integrating eye-
14
15 844 tracking technology. In particular, it should be tested whether it is the stop of the anticipated
16
17 845 target location that causes its detection or whether a slowdown would lead to the same effect
18
19 846 to figure out how closely the polygon is linked to the mechanisms of attentional tracking.
20
21
22

23
24 847 (e) Our findings on gaze behavior in MOT also give rise to the speculation that the use
25
26 848 of overt attention, measured as saccadic reaction times, is affected by spatial uncertainty as
27
28 849 well as by the voluntariness of saccades. First, the ability to use covert attention to detect
29
30 850 changes in the environment can be related to spatial uncertainty because a low contrast in pe-
31
32 851 ripheral stimuli results in higher spatial uncertainty (for an overview, see Carrasco, 2011). In
33
34 852 the current experiment, form changes presumably caused less contrast changes per frame so
35
36 853 that spatial uncertainty can be assumed to be higher in this condition. As a consequence, the
37
38 854 early saccade onset in the form-change condition might result from this uncertainty. In con-
39
40 855 trast, in the motion-change condition, the saccade was initiated later, potentially because of
41
42 856 lower spatial uncertainty. Thus, it seems worthwhile to take spatial uncertainty into account in
43
44 857 future experiments on covert attention in MOT. Furthermore, it can be speculated that sac-
45
46 858 cades were voluntarily controlled. This speculation is substantiated by the fact that saccades
47
48 859 are usually initiated in a bottom-up process resulting in an involuntary overt-attentional shift
49
50 860 as soon as stimuli in the environment change their motion (Drew et al., 2009; Hillstrom &
51
52 861 Yantis, 1994). In our case, those shifts of overt attention were only sparsely observed. Hence,
53
54 862 the assumption seems to make sense that the saccades to the event-target were voluntarily
55
56
57
58
59
60

1
2
3 863 controlled (Peterson, Kramer, & Irwin, 2004) in order to update the target position because of
4
5 864 the higher foveal resolution (Zelinsky & Todor, 2010). Consequently, in a follow-up study, it
6
7 865 is planned to eliminate the task-dependent evocation of a saccade by pursuing a dual task ap-
8
9 866 proach in which a button has to be pressed as soon as a change is detected and, regardless of a
10
11 867 change, all four targets have to be recalled at the end of the trial. If peripheral vision suffi-
12
13 868 ciently allows for change detection, it can be predicted that no or at least less saccades will be
14
15 869 executed because the main task would be to identify the four targets at the end of each trial. If,
16
17 870 by this means, the independence of change detection from subsequent saccades could be con-
18
19 871 firmed, it is also planned to experimentally vary the before discussed uncertainty of the
20
21 872 change event to investigate whether the degree of uncertainty plays a crucial role in the effec-
22
23 873 tiveness of peripheral monitoring in MOT.
24
25
26
27

28 874 In summary, our experimental approach allows for a more ecologically valid examina-
29
30 875 tion on covert attention in MOT. Our empirical results show that the detection of feature
31
32 876 changes depend on visual capabilities like spatial resolution and motion sensitivity, that the
33
34 877 centroid-strategy for MOT seems also to be beneficial for change detection to avoid saccade-
35
36 878 related costs, that feature and spatiotemporal information can be processed with covert atten-
37
38 879 tion by concurrent attentional streams, and that spatial uncertainty might affect the voluntary
39
40 880 control of change detection with either overt or covert attention. In any case, the integration of
41
42 881 eye-tracking methods in the field of attention-related MOT research seems to be worthwhile
43
44 882 as existing findings may appear in a new light and novel research questions may reasonably
45
46 883 be derived. Hence, follow-up questions that have been suggested above are currently investi-
47
48 884 gated in our research group.
49
50
51
52

53 885
54
55
56
57
58
59
60

886 **References**

- 887 Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track? Evidence for a
888 resource-limited attentive tracking mechanism. *Journal of Vision*, 7, 1-10.
- 889 Bahrami, B. (2003). Object property encoding and change blindness in multiple object track-
890 ing. *Visual Cognition*, 10, 949-963.
- 891 Boucart, M., Moroni, C., Thibaut, M., Saffarczyk, S., & Greene, M. (2013). Scene categoriza-
892 tion at large visual eccentricities. *Vision Research*, 86, 35-42.
- 893 Calvo, M. G., & Lang, P. J. (2005). Parafoveal Semantic Processing of Emotional Visual
894 Scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 31,
895 502-519.
- 896 Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, 51, 1484-1525.
- 897 Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention.
898 *Trends in Cognitive Sciences*, 9, 349-354.
- 899 Catteeuw, P., Helsen, W., Gilis, B., Van Roie, E., & Wagemans, J. (2009). Visual scan pat-
900 terns and decision-making skills of expert assistant referees in offside situations. *Journal*
901 *of Sport and Exercise Psychology*, 31, 786-797.
- 902 Davids, K. (1984). The role of peripheral vision in ball games: Some theoretical and practical
903 notions. *Physical Education Review*, 7, 26-40.
- 904 Diamond, M. R., Ross, J., & Morrone, M. C. (2000). Extraretinal control of saccadic suppres-
905 sion. *Journal of Neuroscience*, 20, 3449-3455.
- 906 Drew, T., McCollough, A. W., Horowitz, T. S., & Vogel, E. K. (2009). Attentional enhance-
907 ment during multiple-object tracking. *Psychonomic Bulletin & Review*, 16, 411-417.

- 1
2
3 908 Elfanagely, O., Haladjian, H., Aks, D., Kourtev, H., & Pylyshyn, Z. (2011). Eye-movement
4
5 909 dynamics of object-tracking. *Journal of Vision*, *11*, 1-16.
6
7
8 910 Faubert, J., & Sidebottom, L. (2012). Perceptual-Cognitive Training of Athletes. *Journal of*
9
10 911 *Clinical Sport Psychology*, *6*, 85-102.
11
12
13 912 Fehd, H. M., & Seiffert, A. E. (2008). Eye movements during multiple object tracking: Where
14
15 913 do participants look? *Cognition*, *108*, 201-209.
16
17
18 914 Fehd, H. M., & Seiffert, A. E. (2010). Looking at the center of the targets helps multiple ob-
19
20 915 ject tracking. *Journal of Vision*, *10*, 1-13.
21
22
23 916 Fencsik, D. E., Klieger, S. B., & Horowitz, T. S. (2007). The role of location and motion in-
24
25 917 formation in the tracking and recovery of moving objects. *Perception & Psychophysics*,
26
27 918 *69*, 567-577.
28
29
30 919 Finlay, D. (1982). Motion perception in the peripheral visual field. *Perception*, *11*, 457-462.
31
32
33 920 Franconeri, S. L., Lin, J. Y., Pylyshyn, Z. W., Fisher, B., & Enns, J. T. (2008). Evidence
34
35 921 against a speed limit in multiple-object tracking. *Psychonomic Bulletin & Review*, *15*,
36
37 922 802-808.
38
39
40 923 Goldstein, E. B. (2010). *Sensation and perception* (8th ed.). Belmont, Calif: Wadsworth Cen-
41
42 924 gage Learning.
43
44
45 925 Gralla, V. (2007). *Peripheres Sehen im Sport – Möglichkeiten und Grenzen dargestellt am*
46
47 926 *Beispiel der synchronoptischen Wahrnehmung [Peripheral vision in sports – Possibilities*
48
49 927 *and limits depicted with the example of synchronoptic perception]*. Dissertation, Ruhr
50
51 928 University Bochum.URL: [http://www-brs.ub.ruhr-uni-bochum.de/netahtml/HSS/Diss/](http://www-brs.ub.ruhr-uni-bochum.de/netahtml/HSS/Diss/GrallaVolker/diss.pdf)
52
53 929 [GrallaVolker/diss.pdf](http://www-brs.ub.ruhr-uni-bochum.de/netahtml/HSS/Diss/GrallaVolker/diss.pdf).
54
55
56
57
58
59
60

- 1
2
3 930 Hartmann, E., Lachenmayr, B., & Brettel, H. (1979). The peripheral critical flicker frequency.
4
5 931 *Vision Research*, *19*, 1019-1023.
6
7
8 932 Helsen, W. F., & Starkes, J. L. (1999). A multidimensional approach to skilled perception and
9
10 933 performance in sport. *Applied Cognitive Psychology*, *13*, 1-27.
11
12
13 934 Hillstrom & Yantis (1994). Visual motion and attentional capture. *Perception & Psychophys-*
14
15 935 *ics*, *55*, 399-411.
16
17
18 936 Howe, P. D., & Holcombe, A. O. (2012). Motion information is sometimes used as an aid to
19
20 937 the visual tracking of objects. *Journal of Vision*, *12*, 1–10.
21
22
23 938 Huff, M., & Papenmeier, F. (2013). It is time to integrate: The temporal dynamics of object
24
25 939 motion and texture motion integration in multiple object tracking. *Vision Research*, *76*,
26
27 940 25–30.
28
29
30 941 Huff, M., Papenmeier, F., Jahn, G., & Hesse, F. (2010). Eye movements across viewpoint
31
32 942 changes in multiple object tracking. *Visual Cognition*, *18*, 1368–1391.
33
34
35 943 Iordanescu, L., Grabowecky, M., & Suzuki, S. (2009). Demand-based dynamic distribution of
36
37 944 attention and monitoring of velocities during multiple-object tracking. *Journal of Vision*,
38
39 945 *9*, 2-23.
40
41
42
43 946 Kowler, E. (2011). Eye movements: The past 25 years. *Vision Research*, *51*, 1457–1483.
44
45
46 947 Kredel, R., Klostermann, A., Lienhard, O., Koedijker, J., Michel, K., & Hossner, E. J. (2011)
47
48 948 Perceptual Skill Identification in a Complex Sport Setting. *BIO Web of Conferences*, *1*,
49
50 949 00051. URL: <http://dx.doi.org/10.1051/bioconf/20110100051>.
51
52
53 950 Landry, S. J., Sheridan, T. B., & Yufik, Y. M. (2001). A methodology for studying cognitive
54
55 951 groupings in a target-tracking task. *IEEE Transactions on Intelligent Transportation Sys-*
56
57 952 *tems*, *2*, 92-100.
58
59
60

- 1
2
3 953 Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in
4
5 954 the absence of attention. *Proceedings of the National Academy of Sciences of the United*
6
7 955 *States of America, 99*, 9596-9601.
- 8
9
10 956 Luu, T., & Howe, P. D. (2015). Extrapolation occurs in multiple object tracking when eye
11
12 957 movements are controlled. *Attention, Perception, & Psychophysics*. URL:
13
14 958 <http://link.springer.com/article/10.3758%2Fs13414-015-0891-8>.
- 15
16
17 959 Meyerhoff, H. S., Papenmeier, F., & Huff, M. (2013). Object-based integration of motion
18
19 960 information during attentive tracking. *Perception, 42*, 119–121
- 20
21
22 961 McKee, S. P., & Nakayama, K. (1984). The detection of motion in the peripheral visual field.
23
24 962 *Vision Research, 24*, 25-32.
- 25
26
27 963 Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glis-
28
29 964 sade detection in eye tracking data. *Behavior Research Methods, 42*, 188-204.
- 30
31
32 965 Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an
33
34 966 early vision mechanism independent of higher-order cognition? An individual difference
35
36 967 approach. *Visual Cognition, 11*, 631-671.
- 37
38
39 968 Papenmeier, F., Meyerhoff, H. S., Jahn, G., & Huff, M. (2014). Tracking by location and fea-
40
41 969 tures: Object correspondence across spatiotemporal discontinuities during multiple object
42
43 970 tracking. *Journal of Experimental Psychology: Human Perception and Performance, 40*,
44
45 971 159–171
- 46
47
48 972 Peterson, M. S., Kramer, A. F., & Irwin, D. E. (2004). Covert shifts of attention precede in-
49
50 973 voluntary eye movements. *Perception & Psychophysics, 66*, 398-405.
- 51
52
53 974 Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for
54
55 975 a parallel tracking mechanism. *Spatial Vision, 3*, 179-197.
- 56
57
58
59
60

- 1
2
3 976 Rensink, R. A. (2002) Change detection. *Annual Review of Psychology*, 53, 245-277.
4
5
6 977 Sears, C. R., & Pylyshyn, Z. W. (2000). Multiple object tracking and attentional processing.
7
8 978 *Canadian Journal of Experimental Psychology*, 54, 1-14.
9
10
11 979 St.Clair, R., Huff, M., & Seiffert, A. E. (2010). Conflicting motion information impairs multi-
12
13 980 ple object tracking. *Journal of Vision*, 10, 1–13.
14
15
16 981 Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recogni-
17
18 982 tion: a review. *Journal of Vision*, 11, 1-82.
19
20
21 983 Thorpe, S. J., Gegenfurtner, K. R., Fabre-Thorpe, M., & Bülthoff, H. H. (2001) Detection of
22
23 984 animals in natural images using far peripheral vision. *European Journal of Neuroscience*,
24
25 985 14, 869-876.
26
27
28 986 Williams, A. M., & Davids, K. (1998). Visual search strategy, selective attention, and exper-
29
30 987 tise in soccer. *Research Quarterly for Exercise and Sport*, 69, 111-128.
31
32
33 988 Yantis, S. (1992). Multi element visual tracking: Attention and perceptual organization. *Co-*
34
35 989 *gnitive Psychology*, 24, 295-340.
36
37
38 990 Yu, H. H., Chaplin, T. A., & Rosa, M. G. P. (2015). Representation of central and peripheral
39
40 991 vision in the primate cerebral cortex: Insights from studies of the marmoset brain. *Neuro-*
41
42 992 *science Research*, 93, 47-61.
43
44
45
46 993 Zelinsky, G. J., & Neider, M. (2008). An eye movement analysis of multiple object tracking
47
48 994 in a realistic environment. *Visual Cognition*, 16, 553-566.
49
50
51 995 Zelinsky, G. J., & Todor, A. (2010). The role of “rescue saccades” in tracking objects through
52
53 996 occlusions. *Journal of Vision*, 10, 1-13.
54
55
56
57
58
59
60

1
2
3 997 Zhong, S-H., Ma, Z., Wilson, C., Liu, Y., & Flombaum, J. I. (2014). Why do people appear
4
5 998 not to extrapolate trajectories during multiple object tracking? A computational investiga-
6
7 999 tion. *Journal of Vision*, 14, 1–30.
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Review Only